

周明全 甄逸华 韦颖 著

基于内容图像 检索技术



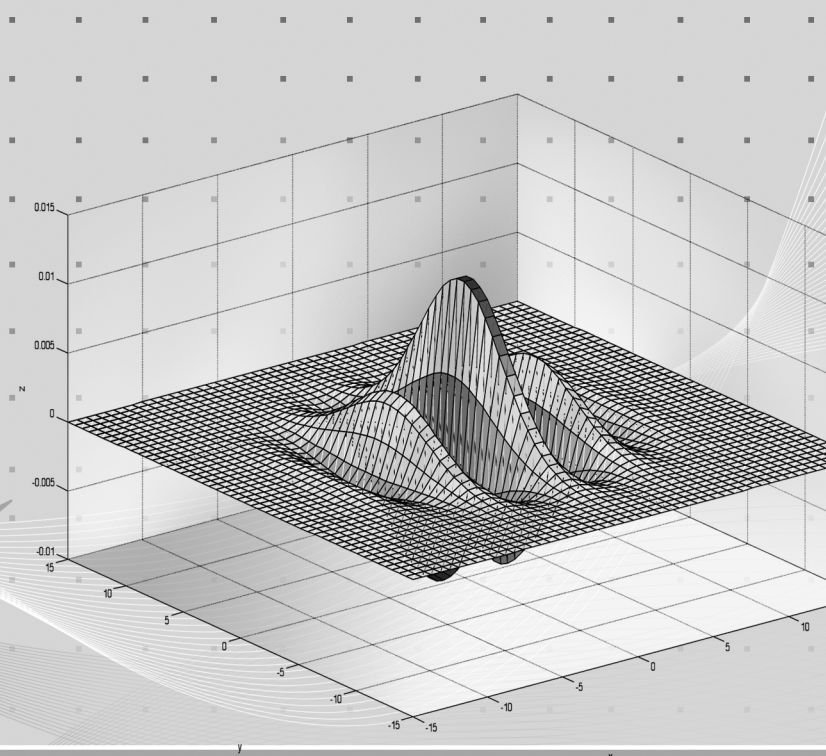
清华大学出版社

基于视觉内容的 图像检索技术研究

*Image Retrieval
Based on the Visual Content*

安志勇
赵 峰
刘 岩
李晋江

编著



電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

基于内容的图像检索技术 (CBIR) 已成为新一代有效的图像数据库检索手段, 是目前国内外广泛关注的研究热点。CBIR 技术涉及计算机视觉、图像处理、图像理解、人工智能、机器学习、统计学、数据库及心理学等诸多领域, 研究内容十分广泛。目前, 国内外针对基于内容的图像检索研究热点较多。本书主要从以下 3 个方面对图像的视觉内容进行探讨和研究: 基于图像视觉内容特征的检索研究, 基于感兴趣区域的检索研究, 基于 Contourlet 变换 (轮廓波变换) 的图像检索技术研究。

本书适合从事图像检索技术 (CBIR) 的科研人员阅读, 也可作为高等学校相关专业研究生的教学用书。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有, 侵权必究。

图书在版编目 (CIP) 数据

基于视觉内容的图像检索技术研究 / 安志勇等编著. —北京: 电子工业出版社, 2011.9

ISBN 978-7-121-14585-8

I. ①基… II. ①安… III. ①图像数据库—情报检索—研究 IV. ①G354.49

中国版本图书馆 CIP 数据核字 (2011) 第 185778 号

责任编辑: 张 剑 (zhang@phei.com.cn)

印 刷:

装 订:

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 720×1 000 1/16 印张: 10.25 字数: 157 千字

印 次: 2011 年 9 月第 1 次印刷

定 价: 58.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

前 言



随着多媒体技术、Internet 技术和大容量存储技术的迅速发展，图像等多媒体数据呈现几何级数的增长。如何有效、快速地从大规模多媒体数据库中检索出所需的数据，是当前具有挑战性的研究课题。图像数据库系统是多媒体信息管理中的重点研究内容，图像数据库技术已广泛应用于数字图书馆、医学、地理信息数据库及国防等众多领域的多媒体信息系统中。目前，传统文本搜索引擎已无法满足大容量的图像数据库检索，因此基于内容的图像检索技术（Content-Based Image Retrieval, CBIR）成为新一代有效的图像数据库检索手段，该技术已成为国内外广泛关注的研究热点。

CBIR 技术涉及计算机视觉、图像处理、图像理解、人工智能、机器学习、统计学、数据库及心理学等多领域的技术，图像内容研究十分广泛。图像的“内容”主要分为两类，即视觉内容和语义内容。视觉内容主要包括颜色、形状和纹理等；而语义内容包括图像的认知和情感，其中图像的认知语义包括对象、场景、对象间关系和情感层特征。目前，国内外针对基于内容的图像检索研究热点较多，本书主要从以下 3 个方面对图像的视觉内容进行探讨和研究。

1) 基于图像视觉内容特征的检索研究 主要包括图像的颜色、形状和纹理特征检索技术。

2) 基于感兴趣区域的检索研究 感兴趣区域的研究属于认知语义中的对象研究，是接近于智能化理解的检索技术，利用图像分割技术获取对象内容，然后针对每个对象的视觉和空间特征等进行基于内容的图

像检索研究。

3) 基于 Contourlet 变换(轮廓波变换)的图像检索技术研究 主要包括近年来利用轮廓波变换进行图像检索的算法研究。

本书共分为 7 章。第一章为绪论,介绍国内外发展现状,以及基于内容的图像检索原理、相似性度量法则、性能评价准则和图像数据库的索引机制等;第二章介绍基于颜色和空间特征的图像检索;第三章介绍基于形状特征的图像检索;第四章介绍基于纹理特征的图像检索;第五章介绍基于感兴趣区域的图像检索;第六章介绍基于兴趣点的图像检索;第七章介绍基于 Contourlet 变换的图像检索。全书由安志勇、赵峰、李晋江和刘岩写作。本书参考了该领域的经典成果和许多其他学者的最新文献,同时也融入了作者近年来在该领域的研究成果。

本书的研究工作得到国家自然科学基金(60970105、60970088)、山东省优秀中青年科学家奖励基金(BS2009DX038、BS2010DX021)、山东省自然科学基金(2009ZRB01620、ZR2009GL011、2009ZRB01737 和 ZR2011FM035)、山东省高校科研发展计划(J09LG22、J10LG21)、山东省信息产业发展专项资金项目(2008X00039)、山东省软科学(2010RKGA2062、2010RKGA1057)、住房城乡建设部项目(2010-K9-54)和烟台市科技发展计划(2011055)等项目的资助。参加研究工作的还有赵珊博士、孙君顶博士、曾智勇博士、崔江涛博士等,在此一并致谢!

另外,要特别感谢范辉教授、原达教授、华臻教授、张鑫教授、谢青松教授、邱勇教授和刘培强副教授,感谢他们为本书的写作提供帮助和支持。

本书可作为从事基于内容的图像检索技术研究和应用的科技人员的参考书,也可以作为高等学校研究生和高年级本科生学习图像检索技术的教学用书。

由于撰写时间和作者水平有限,书中难免存在缺点和错误,敬请读者批评和指正。

编著者

2011 年 6 月 30 日

目 录



第一章 绪论	1
1.1 图像检索的研究背景和意义	1
1.2 图像检索技术的发展现状	3
1.3 图像检索技术的研究内容	5
1.4 图像检索技术基础知识	9
1.4.1 图像内容特征	9
1.4.2 相似性度量准则	11
1.4.3 性能评价准则	15
1.4.4 图像数据库索引机制	17
1.5 本书内容概要	18
第二章 基于颜色和空间特征的图像检索	20
2.1 颜色空间	20
2.1.1 RGB 颜色空间模型	21
2.1.2 HSV 颜色空间模型	21
2.1.3 YCbCr 颜色空间模型	23
2.2 基于颜色特征的图像检索	24
2.2.1 颜色直方图	24
2.2.2 颜色统计矩	26
2.2.3 颜色信息熵	27
2.2.4 颜色视觉关注度	29



2.3 典型的颜色和空间特征的图像检索算法	31
2.3.1 广义颜色直方图	31
2.3.2 颜色集	32
2.3.3 颜色聚合向量	32
2.3.4 颜色相关图	33
2.3.5 SCH 和 geostat 算法	33
2.3.6 局部空间特征	34
2.3.7 基于投影的颜色空间特征	35
2.3.8 环形颜色空间特征	36
2.3.9 广义颜色空间分布矩	38
2.4 本章小结	42
第三章 基于形状特征的图像检索	43
3.1 形状特征概述	43
3.2 基于区域的形状特征分析	44
3.2.1 区域形状参数	44
3.2.2 HU 不变矩	46
3.2.3 Legendre 矩	47
3.2.4 Zernike 矩和 Pseudo-Zernike 矩	48
3.2.5 Tchebichef 矩	50
3.2.6 Krawtchouk 矩	51
3.3 基于边界轮廓的形状分析	52
3.3.1 傅里叶形状描述符	52
3.3.2 基于内角的形状特征	54
3.3.3 小波 HU 不变矩	56
3.3.4 Radon 轮廓不变矩	58
3.3.5 基于显著封闭边界的形状特征	61
3.4 本章小结	64



第四章 基于纹理特征的图像检索	65
4.1 纹理特征概述	65
4.2 空间域的纹理分析	66
4.2.1 Tamura 纹理特征	66
4.2.2 灰度共生矩阵法	68
4.2.3 LBP 算法	69
4.2.4 Laws 纹理能量	71
4.3 频率域纹理分析	73
4.3.1 傅里叶功率谱法	73
4.3.2 小波变换统计矩	74
4.3.3 小波直方图	77
4.3.4 小波 GGD&KLD	79
4.3.5 Gabor 小波纹理分析	82
4.3.6 多小波变换纹理分析	84
4.3.7 双正交小波变换纹理分析	88
4.3.8 复数小波变换纹理分析	90
4.3.9 双密度小波变换纹理分析	93
4.4 旋转不变性纹理特征	95
4.4.1 Polar-wavelet 纹理分析	95
4.4.2 Log-polar 变换纹理分析	97
4.4.3 RIM 纹理分析算法	99
4.5 本章小结	102
第五章 基于感兴趣区域的图像检索	103
5.1 感兴趣区域检索概述	103
5.2 基于视觉显著度模型的图像检索	103
5.3 基于小波感兴趣区域的图像检索	108
5.4 基于颜色和熵的感兴趣区域图像检索	110



5.5 本章小结	113
第六章 基于兴趣点的图像检索	114
6.1 兴趣点概述	114
6.2 兴趣点环形空间特征	117
6.3 兴趣点扇形空间特征	120
6.4 本章小结	122
第七章 基于 Contourlet 变换的图像检索	123
7.1 Contourlet 变换原理	123
7.1.1 拉普拉斯金字塔	124
7.1.2 方向滤波器组	125
7.2 基于 Contourlet 变换的检索算法	127
7.2.1 灰度—平滑共生矩阵	127
7.2.2 广义高斯模型	129
7.2.3 统计矩特征	133
7.3 基于 NSCT 变换的检索算法	133
7.4 圆对称 Contourlet 变换检索	136
7.5 复数 Contourlet 变换检索算法	137
7.6 本章小结	139
参考文献	140

第一章



绪 论

图像是人类表达、记录与交流思想的一种重要手段。随着多媒体技术和大容量数字存储设备技术的迅速发展,图像数据库的数据量呈现几何级数的增长。然而,数字化设备本身并没有提供图像库的管理方法,需要配合特定的软件技术管理和查询图像数据库。因此,如何高效、快速地从海量图像数据库中检索到用户所需要的信息,业已成为当前多媒体数据库领域的一个重要问题。图像数据库的检索研究将对多媒体数字图书馆、医学图像管理、卫星遥感图像、地理信息系统和商标版权管理等方面提供有力的支持。本章主要针对基于内容的图像检索技术基础知识展开论述,主要包括图像检索技术研究内容及发展现状,图像检索的基础知识包括图像内容特征、相似性度量准则、检索算法性能评价、索引机制和查询方式等,以及经典 CBIR 系统介绍。

1.1 图像检索的研究背景和意义

图像数据的急剧增长使得对图像的管理和检索成为关键。早在 20 世纪 70 年代末期,有关学者开始研究图像检索技术。传统的图像检索技术从本质上来说是一种基于文本/数字域结构的图像检索技术,最初的图像检索技术是基于元数据^[1]的检索,每幅图像被当做像素点的简单有序的集合体,用一组元数据的集合来表征,查找时相似性的度量准则



是点对点的比较,通过指定属性来完成图像的查询,其缺点是图像内容较难通过若干个的基本元数据来进行描述,因此较多学者提出了一种基于文本的图像检索技术。基于文本的图像检索技术将图像作为数据库中存储的一个对象,用关键字或自由文本对其进行描述,查询操作是基于该图像的文本描述进行精确匹配或概率匹配。

然而,传统的文本图像检索方法具有下述难以克服的缺点:人工提取注解所需的工作量太大,并且主观性较强;文本描述很难准确地表达图像中丰富的信息内容,不能够充分表达图像的内涵;文本描述常有二义性,往往给准确的检索带来困难。不同的国家和民族很难使用同一语言进行准确加注标志,难以形成统一的检索方法。为了能够方便人们查询和检索多媒体数据,基于内容的图像检索技术(Content-Based Image Retrieval, CBIR)业已成为新一代的有效检索手段。

CBIR 技术的基本思想是根据检索图像所包含的视觉感知内容(包括颜色、纹理、形状等特征)和语义内容建立特征矢量,进而根据特征矢量的相似度进行匹配检索。需要说明的是,图像“内容”的提取不再依赖于传统的人工标注,而是依赖于视觉或语义智能算法。CBIR 属于多媒体的集成综合技术,不仅要确定是否能够找到,而且还要确定相应的输入/输出方法、存储方法、媒体间的组织方法等。CBIR 利用认知科学、用户模型、图像处理、模式识别、知识处理、计算机图形学、数据库管理及信息检索等方法,建立新的媒体数据表示方法和数据模型,采用有效和可靠的查询处理算法,使用户可以在智能化查询接口的辅助下,从多媒体数据库(或大型分布数据库)中,以用户可以接受的响应时间,尽量以与领域无关(在需要时也可以与领域有关)的方式检索到所要求的数据。CBIR 不需要去理解媒体中的对象,关注的是基于内容的快速查找与发现。CBIR 技术具有以下优点。

(1) 基于内容的检索突破了传统的基于关键词检索的局限,直接对图像进行分析并提取特征,使得检索结果更加接近人类视觉感知。



(2) 提取的特征多样化。对图像的特征提取可以是颜色特征、形状特征、纹理特征、轮廓特征、空间关系特征和语义特征等。

(3) 基于内容的图像检索具有很强的交互性，用户可以参与检索过程。

(4) 基于内容的图像检索是一种近似匹配。在检索过程中，采用逐步求精的办法，每次的中间结果是一个集合，不断减小集合的范围，直到定位到目标，与数据库检索的精确匹配算法有明显的不同。

基于内容的图像检索技术的研究将涉及计算机视觉、图像处理、图像理解、人工智能、机器学习、统计学、数据库及心理学等众多领域的技术，可以广泛应用于数字图书馆、防止犯罪、版权保护、医疗图像管理和辅助病情诊断、地理信息系统、卫星遥感图像管理、教育业、网络查询等领域，是一个很有发展前途的研究方向，图像检索技术的研究必将推动相关领域技术的发展。

基于内容的图像检索技术具有广阔的应用前景，因此对该技术的研究具有较为重要的意义，针对这一课题的深入研究将具有重大的理论价值和广泛的应用前景，其成果对我国信息产业的形成与发展将起到积极的促进作用。

1.2 图像检索技术的发展现状

1. 国外图像检索的发展和现状

基于内容的图像检索技术具有广阔的应用前景和巨大的商业价值，因此迅速成为国内外研究的热点。国外关于基于图像内容的检索研究较早，该技术是随着数字图书馆的发展而逐渐发展而成的。1994 年美国启动“数字图书馆”项目。之后，美国国家航空航天局（NASA）、国防部高级研究计划署、国会图书馆、美国国家自然科学基金会、国家医学图书馆等开始研究数字图书馆。随着数字图书馆的迅速发展，基于



内容的图像检索技术成为数字图书馆的关键技术之一。

目前,世界各著名大学、研究机构和公司等已经对图像检索技术进行了深入的研究,其中包括斯坦福大学、麻省理工学院等的研究团队对该技术进行了深入的研究。世界上各大著名的杂志和期刊纷纷设专刊介绍该领域最新的研究成果,如 IEEE Trans. on PAMI、IEEE Trans. on Image Processing、IEEE Trans. on Multimedia、IEEE Trans. on CSVT、Computer Vision and Image Understanding、International Journal of Computer Vision、Pattern Recognition、Signal Processing 等;著名的国际会议如 IEEE Conference on CVPR、IEEE International Conference on ASSP、ACM Conference on Multimedia、SPIE Conference on Electronic Imaging 等纷纷设专题交流最新的研究进展^[2]。

互联网上可以使用的图像检索商业软件系统有 IBM 的 Almaden 研究中心研发的 QBIC^[3]检索系统, MIT 的 Photobook^[4-5]检索系统, Columbia 的 VisualSEEK 和 WebSEEK 检索系统^[6], UCSB 图书馆研发的 Nera 检索系统,除此以外其他的检索系统可以在 <http://www.aalab.cs.uu.nl/cbirsurvey/cbir-survey/cbir-survey.html> 网站上找到,为学习者研究图像检索提供了新的学习平台。这些检索系统具有一定的通用性,然而系统仅使用了颜色、形状和纹理等视觉内容特征,较少使用对象和空间等语义内容特征,在某些情况下检索性能与人类视觉感知不相符合,因此有必要改进和提高这些检索系统的综合性能。

2. 国内图像检索的发展和现状

我国开始于 20 世纪 90 年代后期研究基于内容的图像检索技术,其中科研团队包括中科院各研究所、清华大学和浙江大学等著名研究院所和大学进行了深入研究,相对国外研究团队而言针对该技术的研究较晚。国内有代表性的系统主要有中国科学院计算技术研究所和中国国家图书馆联合开发的“基于特征的多媒体信息检索系统 MIREs”已开始尝试在科学数据库的中药植物数据库、遥感图像检索系统等方面展开应



用,取得了可喜的进展^[7];浙江大学的图像检索研究团队开发了基于颜色的检索系统 Photo Navigator 和基于形状的检索系统 Photo Engine,在此基础上开发了一个基于内容的多媒体检索系统 WebscopeCBR^[8];国防科技大学研发了 MIRC 的多媒体检索系统^[9];清华大学开发了基于 Internet 的基于内容图像检索的原型系统,这些检索原型系统代表了国内的研究水平。

目前,国内基于内容的图像检索技术应用于多个领域:医药卫生领域,利用该技术可以对医学病理图像进行检索,成为辅助医生的诊断和治疗最新手段;工业生产和机械制造领域,可以对企业的产品进行检索和分析,从而提供产品质量的保证;国防科技和军事领域,针对合成孔径雷达(Synthetic Aperture Radar, SAR)图像数据库的检索技术研究可以提高我国的军事力量,SAR 图像具有高分辨、全天时探测能力,突破了传统光学受外界天气的影响,特征信号含有幅度、相位和极化等丰富信息,SAR 图像发展水平的高低已经成为衡量一个国家军事力量的标志之一,因此基于高分辨率的 SAR 图像的技术研究越来越受到其他研究团队的关注^[10];在知识产权领域,图像检索技术应用于专业图像数据库的管理,如商标图像库,可以查询企业的商标信息,从而为企业提供知识产权的保护提供数据。

1.3 图像检索技术的研究内容

一个典型的 CBIR 系统主要由图像库建立子系统和图像检索子系统两部分组成。图像库建立子系统的主要功能是建立图像库,该子系统主要包括图像预处理和特征提取,特征提取对图像库中的图像抽取特定的特征生成特征矢量,并与图像一起存储在图像库中,从而形成基于内容的图像数据库;图像检索子系统的主要功能是负责与用户的交互,当用户提交查询图像后,对它进行分析并提取特征矢量,根据相应的相似性度量准则在图像库中进行匹配,最后根据相似度顺序把查询结果返回



给用户。图 1-1 所示为典型 CBIR 系统的基本结构框架。

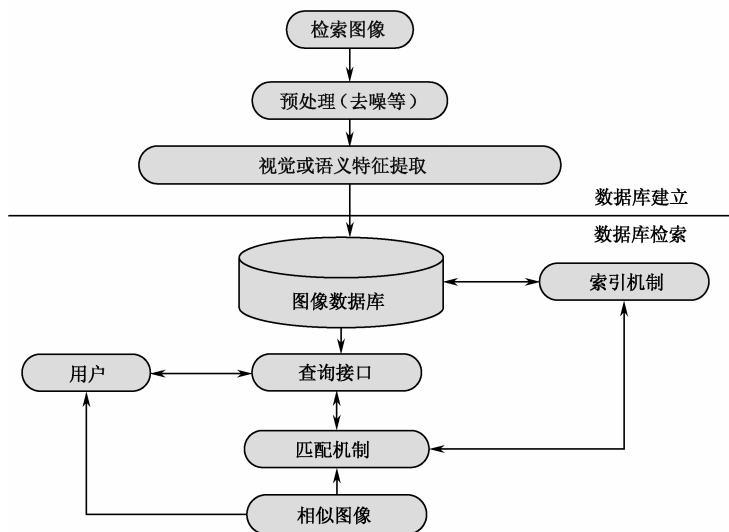


图 1-1 典型 CBIR 系统的基本结构框架

由图 1-1 可知，CBIR 主要涉及 5 个方面的技术，即图像预处理、特征提取、用户查询接口、匹配机制和索引机制。检索过程为，用户通过查询接口查询图像数据库，查询接口根据特征矢量的匹配机制和索引机制将检索结果返回给用户。

1. 图像预处理

图像库建立子系统，需要针对图像提取视觉或语义特征。然而，自然图像由于天气或者拍摄角度等原因造成的图像模糊等情况时，无法正确地提取稳定的图像视觉或语义特征，因此往往需要对图像进行增强或去噪等预处理方法，避免因图像的变形而影响后续的处理。例如，SAR 图像库的检索需要考虑相干斑噪声预处理，相干斑噪声的存在会使图像的信噪比下降，严重时使图像模糊，甚至图像特征消失，因此需要针对 SAR 图像进行相干斑抑制，从而可以提取稳定的图像特征。

2. 特征提取

图像的“内容”特征主要分为两类，即视觉内容和语义内容。视觉



内容特征主要包括颜色、形状和纹理等，这些视觉特征可以利用智能算法自动完成或人机交互获取；而图像的语义特征包含两个层次的特征，即认知层和情感层，其中图像的认知层包括对象、场景、对象间关系，情感层特征主要包括用户对图像内容的理解和个人情感因素，是对图像内容的抽象描述和定性特征，因此需要采用相关反馈或机器学习等人机交互实现。

3. 查询接口

查询接口主要是用户使用相应的关键词查询、草图或示例进行查询，查询接口利用算法对的图像特征向量进行相似性度量，即根据匹配机制将图像进行排序分类，然后将把查询结果按照排序结果返回给用户。因此，查询接口是基于内容图像检索内容中重要的环节，它将涉及人机交互的内容和知识，同时在多特征图像检索过程中涉及特征权重的选取。为了加强系统的检索效果，目前较多学者在查询接口采用相关反馈提高检索性能。

计算机的查询根据查询方式分为要分为关键词查询、草图查询、示例查询和类别查询等，其中草图查询是由用户勾画或编辑的一些元素构成草图，从而基于草图进行查询和检索，检索犯罪嫌疑人信息是草图查询的一种典型应用。

根据图像数据的来源可以分为外部图例查询和内部图例查询。外部图例查询中的示例图像来源于检索系统的外部，可以是 Internet 上的网络图片或其他的图片，该查询可以从图像数据库检索出相似的图像，这种查询的优点是简单易用，使用方便。内部图例查询的示例图片来自于图像数据库，内部图例查询的优点是使用方便。

4. 匹配机制

数据库的匹配机制主要有两种方式，即完全匹配和相似性匹配。完全匹配需要特征向量完全相同，这样才能实现图像匹配，因此属于精确



匹配。相似性匹配是指图像特征向量间的距离测度小于提前设定的阈值时,则图像匹配完成,因此属于模糊匹配。传统的数据库往往是完全匹配,而基于内容的图像检索主要采用相似性匹配,这样用户返回的图像数据往往是按照相似度进行排序。在匹配机制中,重要的是相似性测度距离,相似性测度选择的好坏将直接影响图像检索的性能。

5. 索引机制

索引机制技术是提高检索效率的重要关键技术,是多媒体数据库的研究难点问题。当前,现有的 CBIR 实验原型系统仅处理几百或几千幅图像,采用常用的索引方法如 R-树等,图像数据库的相似度计算不会导致性能效率的显著下降,但这些索引方法在维数足够高的情况下,检索性能下降很快,会出现“维数灾难”现象,这就需要采用高维数据索引方法来加速检索过程^[11]。在传统多维索引技术的基础上探索更加有效的高维索引技术就成为一个急需解决的问题。

上述 5 个部分的研究均属于图像检索的研究范畴。目前,国际上 CBIR 技术的研究热点主要集中在特征提取、查询接口和索引机制部分,主要有以下 6 个方面。

1) 图像视觉内容的特征提取研究 属于特征提取部分,包括颜色、纹理、形状等特征的研究,目前各研究机构和团体已针对 MPEG-7 标准中的视觉特征描述子进行了研究,随着 MPEG-7 标准的完善和规范,会在一定程度上促进图像检索的发展。

2) 基于感兴趣区域的图像检索 属于特征提取部分,感兴趣区域的提取往往需要借助于图像分割算法,即将人类感兴趣的区域或对象提取出来,分析感兴趣区域或对象的空间关系等,然后将针对感兴趣区域进行特征提取,进行相似性度量,并进行图像检索。

3) 基于兴趣点的图像检索 属于特征提取部分,兴趣点往往是图像中最具代表性的视觉关注点,这些点蕴涵了丰富的图像信息,利用兴趣点进行特征提取符合人类的视觉特性,同时其计算复杂度低。



4) 基于高层语义的图像检索 属于特征提取部分, 高层语义包括场景、行为和情感等语义内容的研究, 高层语义内容具有相对主观抽象的特点, 虽然人们偏爱语义查询, 但是这种查询方式的完全智能化还需要进一步深入研究。

5) 高维索引技术的研究 属于高维索引部分, 图像检索的速度和实时性是评价检索算法的一个标准, 然而图像特征维数较高, 检索速度和实时性就成为图像检索的瓶颈, 有效的高维索引技术可以加快检索的速度, 提高检索的实时性。

6) 相关反馈技术的研究 该技术属于查询接口部分, 主要思想是借助人机交互机制, 根据用户的需求动态调整系统检索时所采用的特征向量或参与检索的不同特征的权重系数, 从而尽量缩小底层特征和高层语义之间的差距, 提高算法的检索效果。

目前, 文献[12-16]提出利用相关反馈进行语义传递的工作, 同时利用机器学习与相关反馈相结合进行研究。

1.4 图像检索技术基础知识

1.4.1 图像内容特征

图像检索中常用的视觉内容特征主要包括图像的颜色、形状和纹理等。

1) 颜色特征 颜色是图像的一种重要的视觉特征。由于颜色特征定义比较准确, 抽取相对容易, 所以在图像检索中得到广泛的重视和研究, 并已提出了很多算法。与其他视觉特征相比, 颜色特征对图像的旋转、平移、尺度变化甚至各种形变依赖性比较小, 具有相当强的鲁棒性。为了能用计算机来处理彩色信息, 必须用定量的方法对其描述, 即颜色信息必须在一定的颜色模型下进行表示, 常用的颜色模型主要有 RGB 模型、HSV 模型、HSI 模型和 $CIEL^*a^*b^*$ 和 $CIEL^*u^*v^*$ 模型等^[17]。颜色



基础部分将在第二章重点介绍。

2) 形状特征 形状通常与图像中的特定目标对象有关,包含一定的语义信息。形状特征比颜色特征更高层的特征,因此利用形状来检索可提高检索的准确性和效率。从广义的角度来说,基于形状的检索不仅包括传统意义的二维形状检索,还包括在三维图像的形状检索。通常来说,形状特征有两种表示方法,即轮廓特征和区域特征。图像的轮廓特征只用到物体的外边界,而图像的区域特征则关系到整个形状区域。形状特征适宜于描述对象的几何信息,是目标形状的一种高层描述,但由于图像分割的困难,限制了形状特征的广泛应用。

3) 纹理特征 纹理特征是一种不依赖于颜色或亮度的反映图像中同质现象的视觉特征^[18],因此被广泛应用于模式识别和计算机视觉等领域中。由于纹理分布形态复杂多样,至今人们对纹理的定义没有形成统一的认识^[19]。纹理特征反映了图像灰度模式的空间分布,包含了图像的表面信息和环境的关系,具有确定与随机的二重性、层次性和尺度性等特性,能更好地反映图像的宏观信息和微观结构。

4) 对象特征 当人类观察一幅图像时,往往对图像中的对象(Object)或目标感兴趣,因此基于感兴趣区域的图像检索往往先提取图像的对象目标,然后根据对象的颜色、纹理和形状等视觉内容特征进行图像检索。图像对象的分割往往借助于图像理解和分割算法,根据对象和背景之间的颜色、纹理或边缘等特征的差异,将对象从背景分离出来。基于感兴趣对象的图像检索符合人类以对象为主的思维方式,也可以将检索区域缩小到对象级^[19],但是针对自然图像中没有明确对象目标的图像而言效果不太理想。

5) 空间特征 空间关系是指空间对象之间具有的某种空间度量、方向、拓扑等关系。当图像中包含多个对象目标,对象的空间关系可以弥补其他图像特征空间关系的不足。由于利用图像对象间的空间位置关系来区别图像符合人们识别图像的习惯,而且空间关系特征很容易与其



他的视觉特征相结合,实现多视觉特征的图像综合检索,所以很多研究人员针对基于空间特征的图像检索方法进行了研究。

6) 语义特征 单纯利用颜色等视觉特征进行检索性能不够理想。为了提高检索性能,在检索中往往充分利用图像的语义特征。语义特征包含人类对图像内容的理解,建立视觉内容和语义内容的映射。目前,基于语义图像检索方法分为两大类^[20]:一是自动语义标注,该方法必须利用已标注过的图像数据库作为训练集,采用机器学习或统计学习方法对图像进行标注;二是基于语义空间的图像检索方法,该方法不需要已标注过的图像数据库,而是通过相关反馈的方法创建高层语义空间,进而创建底层特征空间和该语义空间之间的映射,然后将待查图像转变为语义空间中的表示形式,从而实现图像在语义空间中的检索。

1.4.2 相似性度量准则

图像检索往往需要确定待检索图像与图像数据库中图像之间的相似性,因此需要合适的方法表征特征向量间的距离。相似性度量可以确定不同图像之间的相似程度,而这种度量测度往往依赖于特征本身。设 α, β, γ 为任意的特征向量, d 为距离函数,则距离度量函数往往满足四条公理,四条公理描述如下^[2,21]。

➤ 自相似公理

$$d(\alpha, \alpha) = d(\beta, \beta) = 0 \quad (1-1)$$

➤ 最小公理

$$d(\alpha, \beta) \geq d(\alpha, \alpha) = 0 \quad (1-2)$$

➤ 对称公理

$$d(\alpha, \beta) = d(\beta, \alpha) \quad (1-3)$$

➤ 三角不等式公理

$$d(\alpha, \gamma) \leq d(\alpha, \beta) + d(\beta, \gamma) \quad (1-4)$$

下面介绍目前图像检索中常用到的相似性度量方法。设 d 为距离函



数, A 、 B 是两幅图像对应的特征矢量, a_i 、 b_i 代表特征分量。

1. 直方图相交距离

Swain 等学者^[22]针对直方图特征提出直方图相交距离, 算法计算简单快速, 能较好地抑制背景对检索性能的影响, 即

$$d_{hi}(A, B) = 1 - \sum_{i=1}^n \min(a_i, b_i) \quad (1-5)$$

为了降低尺度变化的影响, 归一化后直方图测度数学描述为^[23]

$$d_{hi}(A, B) = 1 - \frac{\sum_{i=1}^n \min(a_i, b_i)}{\min\left(\sum_{i=1}^n a_i, \sum_{i=1}^n b_i\right)} \quad (1-6)$$

2. Minkowsky 范数距离

Minkowsky 范数距离是在范数基础上进行定义的, 即

$$L_p(A, B) = \left[\sum_{i=1}^n |a_i - b_i|^p \right]^{\frac{1}{p}} \quad (1-7)$$

若 $p=1$, $L_1(A, B)$ 则称为城区 (city-block) 距离,

$$L_1(A, B) = \sum_{i=1}^n |a_i - b_i| \quad (1-8)$$

若 $p=2$, $L_2(A, B)$ 则称为欧氏距离 (Euclidean distance),

$$L_2(A, B) = \left(\sum_{i=1}^n (a_i - b_i)^2 \right)^{\frac{1}{2}} \quad (1-9)$$

若 $p \rightarrow \infty$, $L_\infty(A, B)$ 则称为 Chebychv 距离,

$$L_\infty(A, B) = \max_{i=1}^n |a_i - b_i| \quad (1-10)$$

Minkowsky 范数距离是图像检索相似测量经常采用的测度方法, 但这种测度往往不是最优的, 因此还需要根据具体的特征向量设计相似性测度。



3. 相关系数距离

相关系数距离的数学描述如下^[24],

$$d_{\rho} = 1 - \rho(A, B) \quad (1-11)$$

其中相关系数 $\rho(A, B)$ 可定义为

$$\rho(A, B) = \frac{\sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2 \sum_{i=1}^n (b_i - \bar{b})^2}} \quad (1-12)$$

式中, $\bar{a} = \frac{1}{n} \sum_{i=1}^n a_i$, $\bar{b} = \frac{1}{n} \sum_{i=1}^n b_i$ 。

4. χ^2 距离

χ^2 距离的数学描述为^[25]

$$d_{\chi^2}(A, B) = \sum_{i=1}^n \frac{(a_i - M_i)^2}{M_i} \quad (1-13)$$

式中, $M_i = \frac{a_i + b_i}{2}$, χ^2 距离具有特征向量 A, B 的对称性。

5. 二次式距离

二次式距离的数学描述^[26],

$$d_{\text{quad}}(A, B) = (A - B)^T M (A - B) \quad (1-14)$$

式中, $M = [m_{ij}]$, m_{ij} 表征下标为 i 和 j 的两种颜色之间的相似度。颜色相似性矩阵 M 是基于颜色心理学方法进行研究的, 该矩阵不仅分析完全不同颜色之间的相似性度量, 而且分析了相似但不相同的颜色间相似性度量, 因此检索实验表明二次式距离的性能优于欧氏距离和直方图相交距离。

6. KLD 距离 (Kullback-Leibler Distance)

KLD 距离数学描述为^[27],



$$d_{\text{KLD}}(\mathbf{A}, \mathbf{B}) = \sum_{i=1}^n a_i \log \frac{a_i}{b_i} \quad (1-15)$$

式中, $a_i \geq 0$, $b_i \geq 0$ ($1 \leq i \leq n$) 且 $\sum_{i=1}^n a_i = 1$, $\sum_{i=1}^n b_i = 1$ 。KLD 距离对直方图柱值数敏感。由分析可知, 对于 \mathbf{A} , \mathbf{B} 向量而言不具有对称性, 因此效果不够理想。

Minh 提出基于小波变换的广义高斯模型 (Generalized Gaussian Density, GGD) 参数估计的图像检索, 利用 KLD 距离作为相似性度量。子带 G 和 H 的 KLD 距离为^[28]

$$D(p(G; \alpha_1, \beta_1) / p(H; \alpha_2, \beta_2)) = \log \left(\frac{\beta_1 \alpha_2 \Gamma(1/\beta_2)}{\beta_2 \alpha_1 \Gamma(1/\beta_1)} \right) + \left(\frac{\alpha_1}{\alpha_2} \right)^{\beta_2} \frac{\Gamma((\beta_2 + 1)/\beta_1)}{\Gamma(1/\beta_1)} - \frac{1}{\beta_2} \quad (1-16)$$

式中, $p(G; \alpha_1, \beta_1)$ 为子带 G 的密度函数, α_1 , β_1 为子带 G 的广义高斯分布参数估计, $p(H; \alpha_2, \beta_2)$ 为子带 H 的密度函数, α_2 , β_2 为子带 H 的广义高斯分布参数, $\Gamma(x)$ 为伽玛函数。

7. Jeffrey 散度距离

Jeffrey 散度距离数学描述为^[29],

$$d_{\text{jef}}(\mathbf{A}, \mathbf{B}) = \sum_{i=1}^n \left[a_i \log \frac{a_i}{m_i} + b_i \log \frac{b_i}{m_i} \right] \quad (1-17)$$

式中, $m_i = \frac{a_i + b_i}{2}$, Jeffrey 散度距离具有对于 \mathbf{A} , \mathbf{B} 向量而言, 实验证明对高斯噪声具有较强的鲁棒性, 因此效果较为理想。

8. 余弦角距离 (Cosine Distance)

余弦角距离的数学描述为^[30],

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{A}^T \mathbf{B} = |\mathbf{A}| \cdot |\mathbf{B}| \cos \theta \quad (1-18)$$

$$d_{\text{cos}}(\mathbf{A}, \mathbf{B}) = 1 - \cos \theta = 1 - \frac{\mathbf{A}^T \mathbf{B}}{|\mathbf{A}| \cdot |\mathbf{B}|} \quad (1-19)$$



其中, $|A| = (A^T A)^{\frac{1}{2}}$, $|B| = (B^T B)^{\frac{1}{2}}$, 余弦角距离可以计算两个向量间方向距离, 实验表明该距离对离群数据比欧氏距离具有更好的鲁棒性。

9. Canberra距离

Canberra 距离的数学描述为

$$d(A, B) = \sum_{i=1}^n \frac{|a_i - b_i|}{|a_i| + |b_i|} \quad (1-20)$$

式中, $d(A, B)$ 为两个向量 A, B 之间的距离, n 为特征向量的维数, a_i, b_i 则表示向量 A, B 的第 i 个分量。Canberra 距离具有特征向量 A, B 的距离对称性, 同时将特征向量进行了归一化处理, 与欧氏距离相比具有较好的检索性能。

1.4.3 性能评价准则

图像检索算法性能的优劣需要一个客观的性能评价准则, 然而图像检索的性能评价往往具有较强的主观性, 因此目前有一些研究者提出了不同的性能评价准则, 这些评价准则能够有效地对检索效果进行全面评价。下面列举的是 4 种公认的图像检索算法的评价准则^[2,17,31-34]。

1. 命中准确率

设检索图像 Q 所在的相关图像集为 D , 根据检索算法得到 M 个相似图像, 其中属于相关图像集 D 共有 N 幅图像, 则本次算法的命中准确率为

$$P = \frac{N}{M} \times 100\% \quad (1-21)$$

为了降低算法的随机性, 可以将多次检索的平均命中准确率作为该算法准确率。

在实际的检索系统中, 命中准确率与图像数据库有一定的关系, 针对同一算法而言, 不同的图像数据库计算出的命中准确率可能具有较大的差异, 因此还有一些学者提出了其他的性能评价标准。



2. 查准率和查全率

查准率 (Precision) 是指在一次查询过程中, 系统返回的相关图像数目占有所有返回图像数目的比例。查全率 (Recall) 则是指系统返回的查询结果中相关图像数目占图像库中所有相关图像数目 (包括返回的和没有返回) 的比例。查准率和查全率是目前在 CBIR 系统中应用较为广泛的一种评价准则。设 Q 是一幅查询图像, 其查全率和查准率可分别定义为

$$\text{recall} = n / N \quad (1-22)$$

$$\text{precision} = n / T \quad (1-23)$$

式中, N 为主观从图像数据库中找出的与图像 Q 相似的图像数目; n 为一次查询中检索系统检索到的相关图像数目; T 为检索系统自动检索输出的总的图像数目。

一般说来, 查准率和查全率越高越好, 但这两个指标存在相互矛盾, 当查准率较高时, 查全率较低, 反之亦然。因此, 检索系统在这两个指标之间达到一个最优的平衡点就可以。因此, 可以统计多幅查询图像的平均查准率和查全率, 直接分别用它们来衡量图像检索算法的性能, 也可以使用查准率对查全率的曲线来评价算法的性能。另外, 还可以根据查准率和查全率计算出检索效率来评价检索算法, 检索效率可定义为

$$\eta_T = \begin{cases} (n / N) \times 100\%, & N \leq T \\ (n / T) \times 100\%, & N \geq T \end{cases} \quad (1-24)$$

查准率和查全率需要用户在图像库中人工找出与查询图像相似的图像集, 这将耗费大量的人工劳动, 因此这种度量准则对于较小型的图像数据库比较合适。

3. ANMRR 排序评测法

MPEG-7 推荐了 ANMRR (Average Normalized Modified Retrieval Rank) 平均排序评测法。设检索图像 $q_i (i = 1, 2, \dots, n)$, $S(q_i)$ 为图像数



数据库中 与 q_i 相似的图像， $M = \max \{S(q_1), S(q_2), \dots, S(q_n)\}$ ， $K = \min \{4S(q_i), 2M\}$ ， 设与被检索图像相似的图像在结果序列中所处的位置为

$$\text{rank}(k) = \begin{cases} k & k \leq K \\ K+1 & k > K \end{cases} \quad (1-25)$$

则

$$\text{ANMRR} = \frac{1}{n} \sum_{i=1}^n \frac{\sum_{k=1}^{S(q_i)} \frac{\text{rank}(k)}{N(q_i)} - 0.5 - 0.5 \times S(q_i)}{K + 0.5 - 0.5 \times N(q_i)} \quad (1-26)$$

由式 (1-26) 可知， ANMRR 的值越小表明算法的符合人类的视觉感受， 因此检索效果越好。

4. 排序值评测法

设 Q 是一幅查询图像， $F = \{F_1, F_2, \dots, F_n\}$ 为算法检索返回的相似图像集， 设 $\text{rank}(F_i) (i=1, 2, \dots, n)$ 为图像 F_i 在图像序列中的排序值， 则排序值评测法的评价指标为

$$R = \frac{1}{n} \sum_{i=1}^n \text{rank}(F_i) \quad (1-27)$$

$$P = \frac{1}{n} \sum_{i=1}^n \frac{i}{\text{rank}(F_i)} \quad (1-28)$$

在上面两个指标 R 和 P 中， 指标 R 表征本次检索的平均排序值， R 排序值越小表示与查询图像的视觉相似度越高， 因此算法的检索准确率越高； 指标 P 表征本次检索的所有相关图像在靠前列的紧密程度， 因此 P 的取值越大表明本次算法的检索性能越高。

1.4.4 图像数据库索引机制

算法的实时性是评价图像检索的一项性能标准， 然而图像特征维数较高， 检索速度和实时性就成为图像检索的瓶颈， 有效的高维索引技术可以加快检索的速度， 提高检索的实时性。 目前， 一些常见的图像数据



库索引机制主要包括传统的数据库技术中的一些索引方法,如 R-树^[35]、线性四叉树^[35]、SS-树^[36]、GC-树等^[37]均可直接用于图像数据库中,但效果不够理想。

图像高维索引技术包括高维特征的缩减和高维特征的索引两个方面,高维特征缩减的主要文献有 Karhunen-Loeve 变换 (KLT) 和按列聚类^[38-40];图像高维特征索引技术文献有 R 树、聚类技术、使用自组织 (Self-Organization Map, SOM) 神经网络构造树状索引结构的方法等^[11,41-43]。另外,文献[44]在数据库检索中利用三角不等式来减少直接比较的次数,通过建立图像库的代表图像集来进行检索。

1.5 本书内容概要

本书各章节内容安排如下所述。

第一章 绪论。本章介绍了图像检索的研究背景和意义、国内外发展现状、图像检索的研究内容和图像检索技术中的基础知识等,其中基础知识包括图像内容特征、相似性度量法则、性能评价准则和图像数据库的索引机制等。

第二章 基于颜色和空间特征的图像检索。本章介绍了常用的颜色空间,RGB 颜色空间、HSV 颜色空间和 YCbCr 颜色空间;然后介绍了 4 种经典的基于颜色特征的图像检索算法,其中包括颜色直方图、颜色统计矩、颜色信息熵和颜色视觉关注度算法;最后重点介绍了 9 种基于颜色和空间特征的图像检索算法。

第三章 基于形状特征的图像检索。本章介绍了两类形状特征检索算法,即基于区域形状特征和基于边缘形状特征检索算法。基于区域形状特征检索算法包括区域形状参数特征方法、HU 不变矩、Legendre 矩、Zernike 矩、Pseudo-Zernike 矩、Tchebichef 矩和 Krawtchouk 矩等方法。基于边界的方法主要有傅里叶形状描述符、基于内角的形状特征、小波 HU 不变矩、Radon 轮廓不变矩和基于显著封闭边界的形状特征。



第四章 基于纹理特征的图像检索。本章分析了空间域的纹理分析和频率域的纹理分析方法。空间域的纹理分析方法包括 Tamura 纹理特征、灰度共生矩阵法、LBP 算法和 Laws 纹理能量法等。频率域的纹理分析方法主要包括傅里叶功率谱法、小波变换统计矩、小波直方图、小波 GGD&KLD 方法、Gabor 小波纹理分析、多小波变换纹理分析、双正交小波变换纹理分析、复数小波变换和双密度小波变换纹理分析方法。另外，分析了旋转不变的纹理特征分析方法，包括 Polar-wavelet 纹理特征、Log-polar 变换的纹理特征和 RIM 旋转不变纹理分析算法。

第五章 基于感兴趣区域的图像检索。本章介绍了 3 种基于感兴趣区域的图像检索算法，包括基于视觉显著度模型的图像检索、基于小波域感兴趣区域的图像检索，以及基于颜色和熵的感兴趣区域的图像检索算法。

第六章 基于兴趣点的图像检索。本章介绍了常见兴趣点检测算法，同时重点介绍了两种兴趣点检索算法，即兴趣点环形空间特征检索算法和兴趣点扇形空间特征。

第七章 基于 Contourlet 变换的图像检索。本章介绍了 3 种基于传统 Contourlet 变换的图像检索算法，包括灰度—平滑共生矩阵、广义高斯模型和统计矩检索算法，然后介绍了近年来的最新研究成果，包括基于 NSCT 变换的检索算法、圆对称 Contourlet 变换检索算法和复数轮廓波变换检索算法。

第二章



基于颜色和空间特征的图像检索

颜色是一种重要的图像视觉特征，颜色往往与图像中所包含的物体或场景十分相关，并且颜色特征具有与生俱来的旋转、平移、方向和尺度变化不变性，甚至对各种变形都不敏感，表现出相当强的鲁棒性。因此，利用颜色特征进行图像检索的技术受到了普遍的重视，国内外许多学者对此进行了研究，基于颜色特征的检索方法被广泛应用于基于内容的图像检索中。

2.1 颜色空间

很多学者研究发现，白光是由不同颜色的光线相混合而组成的，而这些不同颜色的光线实际上是不同频率的电磁波，波长从 350nm 到 780nm，人的脑、眼将不同频率的电磁波感知为不同的颜色。神经生理学实验发现，在视网膜上存在 3 种不同的颜色感知器即三原色：红色（R，Red），绿色（G，Green），蓝色（B，Blue），它们相应的波长分别为 700nm、546.1nm、435.8nm，所有颜色都可以看做是 3 个基本颜色的不同组合^[45]。颜色模型是颜色空间表示的基础，为颜色量化和特征提取提供重要的基础保证。在实际应用中常用的颜色模型很多，在图像检索中，主要采用的颜色空间模型有 RGB、HSV 和 YCrCb 等。



2.1.1 RGB颜色空间模型

RGB 颜色模型是一种面向硬件设备的颜色空间模型，数字图像处理中的其他颜色模型往往需要从 RGB 颜色模型进行转换。RGB 颜色空间模型包括 R 、 G 、 B 三个坐标轴，如图 2-1 所示。

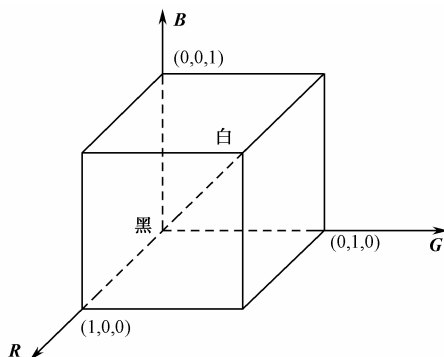


图 2-1 RGB 颜色空间模型

由图 2-1 可知，原点对应黑色，离原点最远的顶点对应白色，立方体与 3 个坐标轴的交点对应于三原色，即红色、绿色和蓝色，剩余的 3 个顶点对应于三补色，即品红（即红加蓝）、蓝绿（即绿加蓝）和黄（即红加绿）。在这个模型中，从黑到白的灰度值分布在从原点到离原点最远顶点间的连线上，而立方体内其余各点对应的不同颜色可以用该点到原点的向量来表示。因此，每幅彩色图像都可以分解到 3 个原色平面上进行表示，从而为进一步的颜色处理提供基础。

RGB 颜色空间模型一般用于计算机图形学和图像显示，是一种面向硬件设备的颜色空间^[17]。

2.1.2 HSV颜色空间模型

RGB 颜色空间模型不符合人类视觉感知，因此在图像检索系统中常用 HSV 颜色空间模型。HSV 颜色空间模型是一种面向视觉感知的模型，通过色调（Hue）、饱和度（Saturation）和亮度（Value）来表示图



像颜色。色调是物体反射的光线中主要占优光波的波长决定的，不同波长产生不同颜色的感觉，就是指色调不同；饱和度是指色彩的深浅程度，饱和度高表示颜色深，而饱和度低则表示颜色浅；亮度是指色彩的明亮程度，需要注意的是它与光强度没有直接的关系。图 2-2 所示的是 HSV 颜色空间模型。

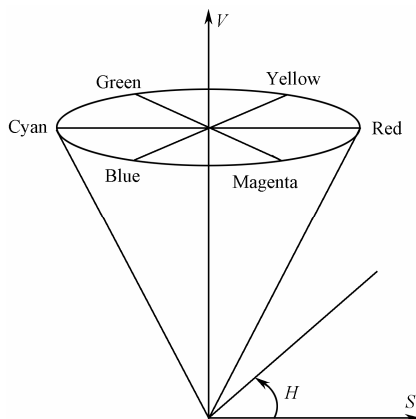


图 2-2 HSV 颜色空间模型

通常 RGB 颜色空间模型转换为 HSV 模型的计算公式为^[46]

$$H = \begin{cases} \theta & G \geq B \\ 2\pi - \theta & G < B \end{cases}, \text{ 其中 } \theta = \cos^{-1} \left[\frac{\frac{1}{2}[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right] \quad (2-1)$$

$$S = 1 - \frac{\sqrt{3}}{V} \min(R, G, B) \quad (2-2)$$

$$V = \frac{1}{\sqrt{3}} [R + G + B] \quad (2-3)$$

从 HSV 颜色模型到 RGB 颜色模型的转换公式如下所述^[46]。

1) 当 $H \in [0^\circ, 120^\circ)$ 时,

$$R = \frac{V}{\sqrt{3}} \left[1 + \frac{S \cos(H)}{\cos(60^\circ - H)} \right], \quad B = \frac{V}{\sqrt{3}} (1 - S), \quad G = \sqrt{3}V - R - B \quad (2-4)$$



2) 当 $H \in [120^\circ, 240^\circ)$ 时,

$$G = \frac{V}{\sqrt{3}} \left[1 + \frac{S \cos(H - 120^\circ)}{\cos(180^\circ - H)} \right], \quad R = \frac{V}{\sqrt{3}} (1 - S), \quad B = \sqrt{3}V - G - R \quad (2-5)$$

3) 当 $H \in [240^\circ, 360^\circ)$ 时,

$$R = \frac{V}{\sqrt{3}} \left[1 + \frac{S \cos(H - 240^\circ)}{\cos(300^\circ - H)} \right], \quad B = \frac{V}{\sqrt{3}} (1 - S), \quad G = \sqrt{3}V - R - B \quad (2-6)$$

2.1.3 YCbCr 颜色空间模型

YCbCr 模型是 DVD、摄像机、数字电视等消费类视频产品中常用的颜色空间, 其中 Y 表示亮度分量, Cb 表示蓝色色度分量, 而 Cr 表示红色色度分量。人的肉眼对视频的 Y 分量更敏感, 因此在通过对色度分量进行子采样来减少色度分量后, 肉眼将察觉不到的图像质量的变化。由于人眼视觉系统对颜色的敏感程度低于亮度的敏感程度, 因此 Cb 和 Cr 分量可以采用比 Y 更低的分辨率来表示, 这样在视觉质量要求不高的环境中, 可以减少表示色度分量的数据量, 从而减少存储或传输的数据量需求。YCbCr 颜色空间也是一种用于 JPEG 数字图像的颜色标准^[47], 一般而言, 由 RGB 模型转换为 YCbCr 模型的计算公式为^[19]

$$Y = k_R R + (1 - k_B - k_R)G + k_B B \quad (2-7)$$

$$Cb = \frac{0.5}{1 - k_B} (B - Y) \quad (2-8)$$

$$Cr = \frac{0.5}{1 - k_R} (R - Y) \quad (2-9)$$

由 YCbCr 模型转换为 RGB 模型的计算公式如下^[19],

$$R = Y + \frac{1 - k_R}{0.5} Cr \quad (2-10)$$

$$G = Y - \frac{2k_B(1 - k_B)}{1 - k_B - k_R} Cb - \frac{2k_R(1 - k_R)}{1 - k_B - k_R} Cr \quad (2-11)$$

$$B = Y + \frac{1 - k_B}{0.5} Cb \quad (2-12)$$



国际电信联盟无线通信部 (ITU-R) 推荐 BT.601 中定义 $k_B = 0.114$ 和 $k_R = 0.299$ 。

2.2 基于颜色特征的图像检索

2.2.1 颜色直方图

Swain^[18]首先提出了颜色直方图特征, 颜色直方图反映了图像中颜色的组成分布, 该特征对于模糊、旋转、平移和尺度等具有不敏感性。设图像 f 的大小为 $M \times N$, $f(x, y)$ 表示像素点 (x, y) 处的颜色值, 图像所包含的颜色集记为 C , 则图像的颜色直方图为

$$h_c = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \delta(f(i, j) - c), \quad \forall c \in C \quad (2-13)$$

颜色直方图适合于检索图像的全局颜色相似性的场合, 即通过比较颜色直方图的差异来衡量两幅图像在颜色全局分布上的差异。根据色彩的不同范围和主观颜色感知进行量化, 将色调量化为 8 个级别, 饱和度和亮度分别量化为 3 个级别, 将 HSV 空间划分为 $8 \times 3 \times 3 = 72$ 个颜色区间, 可以将 3 个分量合并为一维特征向量 $C = 9 \times H + 3 \times S + V$, 则 C 的取值范围为 $[0, 71]$, 量化方式为

$$s = \begin{cases} 0 & s \in [0, 0.2] \\ 1 & s \in [0.2, 0.7] \\ 2 & s \in [0.7, 1] \end{cases} \quad (2-14)$$

$$V = \begin{cases} 0 & v \in [0, 0.2] \\ 1 & v \in [0.2, 0.7] \\ 2 & v \in [0.7, 1] \end{cases} \quad (2-15)$$

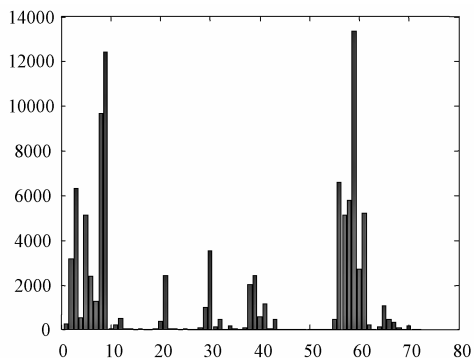


$$H = \begin{cases} 0 & h \in [316, 20] \\ 1 & h \in [21, 40] \\ 2 & h \in [41, 75] \\ 3 & h \in [76, 155] \\ 4 & h \in [156, 190] \\ 5 & h \in [191, 270] \\ 6 & h \in [271, 295] \\ 7 & h \in [296, 315] \end{cases} \quad (2-16)$$

根据式 (2-14)、式 (2-15) 和式 (2-16) 就可以将彩色图像划分为 72 个颜色区间, 进一步就可以进行颜色直方图特征提取。图 2-3 所示为公交车图像的颜色直方图。



(a) 公交车示例图



(b) 颜色直方图

图 2-3 公交车图像的颜色直方图

然而颜色直方图存在特征维数过高和丢失了颜色的空间分布信息等缺点, 在一定程度上容易造成误检现象, 因此较多改进的算法被提出。Stricker^[48]提出累加颜色直方图方法, 统计直方图的零值对相似性度量存在一定影响, 累加颜色直方图方法可以使相似性度量正确反映图像之间的颜色差别, 同时也克服了一般直方图量化过细、过粗检索效果都会下降的缺陷。文献[49-51]等提出利用色调直方图来表示图像颜色特征。文献[52]提出了广义直方图的检索方法, 首先计算出原图像在 4 个方向上的平滑图像, 并将其分别与原图像对应的像素点组成二元组, 统计每



个二元组出现的概率，构造出 4 个方向的广义直方图，最后综合 4 个广义直方图进行检索，广义直方图同时考虑了像素间的方向和局部细节信息。文献[53-55]等提出将模糊理论引入到图像检索系统中，利用模糊颜色直方图提高彩色图像的检索精度。

2.2.2 颜色统计矩

由于颜色分布信息主要集中在低阶矩，Stricker 等^[48]在颜色直方图特征的基础上提出利用一阶矩、二阶中心矩和三阶中心矩表征图像的颜色特征。与颜色直方图相比，该方法在一定程度上降低了颜色特征的维数，可以提高检索的实时性。一阶中心矩表征了邻域颜色的“平均值”，而二阶中心矩和三阶中心矩则表征了邻域颜色与“平均值”的偏差程度。设 h_{ij} 表示第 i 个颜色通道分量中灰度为 j 的像素出现的概率， n 表示灰度级数，颜色矩的数学表达形式为

$$\mu_i = \frac{1}{n} \sum_{j=1}^n h_{ij} \quad (2-17)$$

$$\sigma_i = \left(\frac{1}{n} \sum_{j=1}^n (h_{ij} - \mu_i)^2 \right)^{\frac{1}{2}} \quad (2-18)$$

$$s_i = \left(\frac{1}{n} \sum_{j=1}^n (h_{ij} - \mu_i)^3 \right)^{\frac{1}{3}} \quad (2-19)$$

颜色矩特征表征简单，但实验发现图像低阶矩的检索效率比颜色直方图的检索效率要低，因此在实际检索应用中往往将颜色矩同其他图像特征联合应用。为了提高颜色矩的检索性能，黄朝兵^[56]提出多统计矩颜色直方图特征方法，首先计算彩色图像每个特征点的 3×3 邻域、 5×5 邻域和 7×7 邻域的统计矩，然后计算这些特征点统计矩的直方图分布。设坐标为 (i, j) 的点像素的颜色特征为 $C(i, j)$ ，则其 3×3 邻域的一阶矩、二阶中心矩分别为



$$\begin{cases} m_3(i, j) = \frac{1}{9} \sum_{x=i-1}^{i+1} \sum_{y=j-1}^{j+1} C(x, y) \\ \sigma_3(i, j) = \left[\frac{1}{9} \sum_{x=i-1}^{i+1} \sum_{y=j-1}^{j+1} (C(x, y) - m_3(i, j))^2 \right]^{1/2} \end{cases} \quad (2-20)$$

其 5×5 邻域的一阶矩、二阶中心矩分别为

$$\begin{cases} m_5(i, j) = \frac{1}{25} \sum_{x=i-2}^{i+2} \sum_{y=j-2}^{j+2} C(x, y) \\ \sigma_5(i, j) = \left[\frac{1}{25} \sum_{x=i-2}^{i+2} \sum_{y=j-2}^{j+2} (C(x, y) - m_5(i, j))^2 \right]^{1/2} \end{cases} \quad (2-21)$$

其 7×7 邻域的一阶矩、二阶中心矩分别为

$$\begin{cases} m_7(i, j) = \frac{1}{49} \sum_{x=i-3}^{i+3} \sum_{y=j-3}^{j+3} C(x, y) \\ \sigma_7(i, j) = \left[\frac{1}{49} \sum_{x=i-3}^{i+3} \sum_{y=j-3}^{j+3} (C(x, y) - m_7(i, j))^2 \right]^{1/2} \end{cases} \quad (2-22)$$

对于给定一幅 $n \times n$ 大小的图像, 令 $Q = \{0, 1, 2, 3, \dots, 71\}$, 上述各邻域一阶矩、二阶中心矩经过四舍五入取整后均在 Q 的范围内, 因此就可以分别统计这 6 个特征向量的直方图, 进一步就可以进行相似性度量。然而, 多统计矩颜色直方图的特征维数较高, 因此检索的实时性不强。

2.2.3 颜色信息熵

为了克服直方图的维数较高的问题, Zachary^[57]提出采用图像颜色的信息熵表示图像的颜色特征, 从而可以将特征维数进行降维, 提高检索的实时性。设归一化后的颜色直方图表示为 (h_1, h_2, \dots, h_n) , 若将颜色直方图看做是颜色在图像空间中的概率密度函数, 则图像的信息熵可表示为

$$E(H) = - \sum_{c=1}^n h_c \log_2(h_c), \forall c \in C \quad (2-23)$$



颜色信息熵可以有效降低图像的特征维数,但单纯使用信息熵作为图像的特征进行检索时,其检索性能不够理想,因此可以将信息熵和其他图像特征(颜色、纹理和空间关系等)综合进行检索,从而可以提高检索性能。

另外,Shannon 信息熵具有对称性、非负性、极值性、扩展性和确定性,其中对称性会在一定程度影响图像检索的效果,即对于具有不同视觉特征的直方图,如果它们具有相同的概率分布,则会具有相同的信息熵。孙君顶等^[58]对此问题进行了研究,提出了直方图排序法、直方图面积法、直方图排序与面积法的线性组合法 3 种方法进行提高图像检索性能。

1. 直方图排序法

直方图排序法是在计算直方图信息熵时,引入了加权函数 $f_1(x)$,

$$f_1(x) = 1 + \frac{m_x}{m_{\max}} \quad (2-24)$$

式中, x 表示任意的颜色直方图; m_x 表示直方图 x 经过排序后,各分量移动的总次数; m_{\max} 表示排序时直方图分量需要移动的最大次数,即对逆序直方图进行排序时直方图各分量移动的次数。从而图像信息熵可表示为

$$E(H) = f_1(H) \sum_{i=1}^n h_i \log(h_i) \quad (2-25)$$

2. 直方图面积法

直方图面积法是采用直方图面积消除熵对称性的影响,其原理是对于具有相同或相近概率分布的直方图来说,如果直方图中各分量的次序不同,直方图所对应的面积一般也是不同的。为此,在计算直方图的信息熵时,可引入了加权函数 $f_2(x)$,

$$f_2(x) = 1 + \frac{A_x}{A_{\max}} \quad (2-26)$$

式中, x 为任意图像的直方图特征; A_x 为直方图 x 对应的面积; A_{\max} 为



直方图最大面积。利用直方图面积法比直方图排序法取得更好的效果，由于面积法既考虑了直方图中各分量的位置因素，又考虑了该位置的分量所对应的概率大小。

3. 直方图排序法与面积法的线性组合法

虽然采用直方图面积法的改进效果总体上优于直方图排序法，但对于一些特殊情况，直方图面积法的改进效果反而比不上直方图排序法。为此，可以进一步采用直方图排序法和直方图面积法的线性组合来进一步增强图像的检索效果。采用两种方法线性组合的权函数可表示为

$$f(x) = r_1 f_1(x) + r_2 f_2(x) \quad (2-27)$$

式中， r_1 ， r_2 为权值系数，且 $r_1 + r_2 = 1, r_1, r_2 \in [0, 1]$ 。该方法有效地综合了直方图排序法和面积法的优点，具有较好的检索性能。

2.2.4 颜色视觉关注度

沈云涛^[59]根据人类的视觉感知系统提出了颜色视觉关注度特征，通过提升视觉显著度高的颜色相应特征的关注度权值，补偿由于感知识别系统的颜色敏感度随颜色变化而引起的视觉差异，实验证明该算法切实有效。人类的视觉感知系统对不同颜色显示出不同的分辨力和关注度，为了使颜色特征能够更准确地描述视觉系统的主观感受，引入颜色的关注度函数 $f(h, s, v)$ 。假设视觉系统对黑色的关注度为 1，彩色的视觉关注度相对于黑色的增量由两部分组成：亮度值增大引起的关注度增量 $f_v(v)$ ，色度值变化引起的关注度增量 $f_c(h, s, v)$ ，一般形式的颜色关注度函数 $f(h, s, v)$ 为

$$f(h, s, v) = 1 + f_v(v) + f_c(h, s, v) \quad (2-28)$$

感知系统作为非线性系统对外界刺激量的响应幅值并不与外界刺激量成正比，因此根据 Stevens 法则 $f_v(v)$ 可定义为

$$f_v(v) = k_v \cdot v^{0.6} \quad (2-29)$$

式中， k_v 为参数。图 2-4 所示为 Stevens 法则示意图。

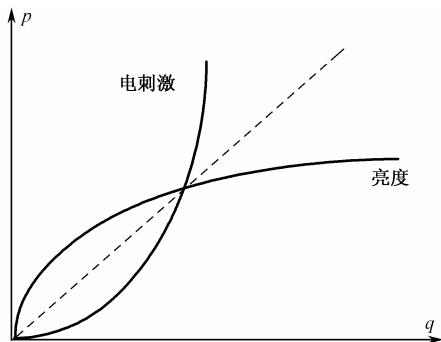


图 2-4 Stevens 法则示意图

相对色调而言，视觉系统对颜色饱和度的感知较为简单，并且饱和度感知与亮度感知比较相似，它的非线性程度要低于亮度感知的非线性， $f_s(s)$ 可定义为

$$f_s(s) = k_s \cdot s^t \quad (2-30)$$

式中， k_s 为参数， $0 < t < 1$ 。根据色调对颜色关注度的影响，色调关注度函数 $f_h(h)$ 可以定义为

$$f_h(h) = \frac{\cos(\text{mod}(h, \frac{\pi}{3}) - \frac{\pi}{6})}{\cos(\frac{\pi}{6})} \quad (2-31)$$

色度感知是相应颜色色调、饱和度和亮度共同作用的结果。量化色度变化对关注度的影响，必须考虑颜色的亮度因素，因此 $f_v(h, s, v)$ 描述色度的变化的视觉关注度增量为

$$f_c(h, s, v) = k_c \cdot f_v(v) \cdot f_s(s) \cdot f_h(h) \quad (2-32)$$

将式 (2-29)、式 (2-30)、式 (2-31) 和式 (2-32) 代入到式 (2-28)，即可得到颜色关注度函数 $f(h, s, v)$ 。针对每幅图像的颜色关注度，就可以研究计算颜色关注度的直方图特征，从而可以进行图像检索。颜色关注度的直方图与传统直方图存在如下区别。

(1) 后者在计算落入某颜色量化区间内图像像素点的个数时，并未区分不同颜色点间的信息量的差别，每点均以 1 记入个数；而前者考虑



到颜色的视觉差异性,以相应颜色的关注度记入个数。

(2)颜色视觉关注度表示某一种颜色中所占视知觉系统感受到的总信息量的百分比。

在实际的应用系统中,将颜色视觉关注度直方图 and 传统颜色直方图融合进行图像检索,具有较高的检索性能。

2.3 典型的颜色和空间特征的图像检索算法

颜色直方图、颜色矩、颜色熵 and 颜色视觉关注度方法所描述的颜色特征是图像的全局颜色特征,不包括图像颜色的空间分布特征,仅利用这些特征进行图像检索时极易造成误检现象。因此,基于颜色和空间特征的图像检索算法成为研究热点。下面是 9 种典型的颜色和空间特征的图像检索算法。

2.3.1 广义颜色直方图

传统直方图方法不能描述颜色在图像的空间分布,文献[51,60]提出了广义颜色直方图方法。广义颜色直方图是在广义图像的基础上进行提取的,设彩色图像 $f(x,y)$ 的大小为 $M \times N$,将 $f(x,y)$ 图像经过 3×3 、 5×5 或 7×7 点阵平滑后的图像为 $g(x,y)$,由 $f(x,y)$ and $g(x,y)$ 构成一个二元组 (f,g) ,即为图像 $f(x,y)$ 的广义图像。

张海成等[51]根据色彩 and 主观颜色感知量化为 72 个颜色区间,因此广义图像的 $f(x,y)$ and $g(x,y)$ 两个分量均被划分为 72 个颜色区间,进一步可以分析图像的颜色特征。针对广义图像的每个特征点的颜色进行统计,就得到了二维图像直方图 $H = [h(0,0), h(0,1), \dots, h(i,j), \dots, h(71,71)]$,其中 $i, j \in [0, 71]$ 。

广义直方图考虑了图像中相邻像素之间的相关性,即将图像的颜色信息 and 空间信息进行结合,可以有效提高图像检索的性能。然而,广义直方图的维数较高,张海成提取了图像的主色调广义直方图。主色调广



义直方图既考虑了图像颜色信息和空间信息的结合，又在 HSV 空间中有效地压缩了颜色矢量的维数，节省了存储空间，提高了检索性能。由于广义直方图具有较好的降噪能力，目前已广泛应用于图像增强等领域中^[61]。

2.3.2 颜色集

颜色集^[62]首先将 RGB 颜色空间转化成视觉均衡的颜色空间（如 HSV 空间），并将颜色空间量化成若干个 bin，然后利用颜色自动分割技术将图像分为若干区域，每个区域用量化颜色空间的某个颜色分量来索引，从而将图像表达为一个二进制的颜色索引集。颜色集是颜色直方图的一种近似，相似性度量可以比较不同颜色集和区域之间的空间距离。颜色集特征实质上是一种二进制的特征向量，可以利用二分查找树来加快检索速度，提高系统检索性能。

2.3.3 颜色聚合向量

Pass 等^[63]针对直方图算法的改进，提出了颜色聚合向量的方法。颜色聚合向量将直方图中每一个颜色簇划分成聚合的和非聚合的两部分。在图像相似性比较过程中，分别比较它们的相似性，然后综合权衡后得到一个相似值，从而得出结果。彩色图像经过量化后，就可以通过计算连通域将图像中的像素分为聚合和非聚合两类，判断某一连通区域是否为聚合时，阈值一般被设为某一固定值，大于等于这一固定值，则判为聚合，否则为非聚合。假设 α_i 和 β_i 分别代表直方图的第 i 个区间中聚合像素和非聚合像素的数目，图像的颜色聚合向量可表示为

$$\langle (\alpha_1, \beta_1), (\alpha_2, \beta_2), \dots, (\alpha_N, \beta_N) \rangle \quad (2-33)$$

而 $\langle \alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_N + \beta_N \rangle$ 就是该图像的颜色直方图。颜色聚合向量考虑图像颜色的空间分布信息，具有较好的检索性能。



2.3.4 颜色相关图

Jing Huang^[64]提出颜色相关图算法,传统的颜色直方图仅刻画了某一种颜色的像素数目占像素总数目的比例,只是一种全局的统计关系,而颜色相关图则表达了颜色随距离变换的空间关系,也就是说,颜色相关图不仅包含图像颜色统计信息,同时包括颜色之间的空间关系。对于任意的图像 I ,假设 $I_{c(i)}$ 表示图像中颜色为 $c(i)$ 的所有像素的集合,则颜色相关图可以表示为

$$r_{i,j}^{(k)} = \Pr_{p_1 \in I_{c(i)}, p_2 \in I} [p_2 \in I_{c(j)}, |p_1 - p_2| = k \mid p_1 \in I_{c(i)}] \quad (2-34)$$

式中, $i, j \in \{1, 2, \dots, n\}$, n 表示图像颜色级数, $k \in \{1, 2, \dots, d\}$, d 表示在计算图像的相关图时所设定的像素间的最大距离, $|p_1 - p_2|$ 表示像素 p_1 和 p_2 间的距离。

一般而言,颜色相关图所需的空间复杂度为 $O(n^2 d)$,复杂度较高,为此可以考虑图像空间中具有相同颜色像素间的空间关系颜色自相关图 (Color auto-correlogram),这样空间复杂度降为 $O(nd)$,可以在一定程度上提高检索性能。

近年来,孙君顶^[65]等在颜色相关图的基础上提出状态关联图 (State Correlogram) 来反映状态空间中不同的状态点按照空间距离的分布特征,图像的状态关联矩阵不仅反映了图像的空间位置关系,也反映了图像的形状信息。实验表明,该方法的检索结果能较好地接近人的视觉感知效果。

2.3.5 SCH 和 geostat 算法

Clique^[66]等人提出采用 (Spatial-Chromatic Histogram, SCH) 来描述具有相同颜色的像素在图像空间的分布特征。设 $C_i = (x_i, y_i)$ 为图像颜色 i 的所有像素的质心,设 σ_i 表示颜色为 i 的像素同其质心 C_i 之间距离的标准差,其定义为



$$\begin{cases} x_i = \frac{1}{|A_i|} \sum_{(x,y) \in H_i} x_i, y_i = \frac{1}{|A_i|} \sum_{(x,y) \in A_i} y_i \\ d(p, C_i) = \sqrt{(x - x_i)^2 + (y - y_i)^2} \\ \sigma_i = \sqrt{\frac{\sum d(p, C_i)^2}{|A_i|}} \end{cases} \quad (2-35)$$

式中 p 表示 (x, y) 处的像素。对于任意一幅图像 q ，其 SCH 索引特征可表示为， $(h_1^q, \sigma_1^q, \dots, h_i^q, \sigma_i^q, \dots, h_n^q, \sigma_n^q)$ ，其中 h_i^q 表示图像 q 中颜色 i 的直方图特征，而 σ_i^q 反映了图像中颜色为 i 像素的空间分布特征， n 为图像 q 中颜色的数目。

SCH 中的取值受到图像中像素数目的影响，因此 SCH 索引特征受到图像尺寸大小的影响，不具备尺度不变性，为此，Lim^[67]等人提出了 geostat 算法，利用地理信息统计特性的松散参数 (Looseness Parameter) 来描述图像颜色的空间分布特征，其定义为

$$\begin{cases} L_i = \frac{1}{|A_i|} 2\pi V_i \\ V_i = \frac{1}{|A_i|} \sum d(p, C_i)^2 \end{cases} \quad (2-36)$$

其索引特征可以表示为 $(h_1^q, L_1^q, \dots, h_i^q, L_i^q, \dots, h_n^q, L_n^q)$ ，其中 h_i^q 表示图像 q 中颜色 i 的直方图特征，而 L_i^q 反映了图像中颜色为 i 像素的空间分布特征， n 为图像 q 中颜色的数目。由于对提取的距离 $d(p, C_i)$ 进行归一化处理，因此 geostat 方法提取的分布特征不仅具有平移不变性、旋转不变性，而且具有尺度不变性。基于自然图像的检索实验表明，geostat 方法在一定程度上优于 SCH 算法。

2.3.6 局部空间特征

为了获取图像的局部空间颜色特征，文献[68-69]提出将图像进行分割或分块，提取图像块的颜色特征作为空间分布，具有较好的检索效果。



Hsu 等^[70]将图像划分为矩形区域,提出将每个矩形区域的一种主要颜色作为空间特征。王涛^[69]提出将图像内容看做是由若干对象组成的集合,利用图像分割得到主要对象,然后根据对象的颜色、位置和形状特征计算图像间内容的相似度。实验表明,当图像中有明显的物体时,该方法与颜色直方图相比,能够更加准确和高效地查找出用户所需内容的图像,明显地提高了检索精度。

汪慧兰等^[68]提出基于图像中主要物体的尺寸动态地对图像分块来获得颜色的空间分布特征,并选取了符合人眼感知特性的 HSV 颜色空间进行非均匀量化,提取各个分块的主色及其所占的百分比作为颜色综合特征矢量,根据所提出的综合特征矢量给出相应的相似度计算方法,同时构造出相似矩阵进行基于颜色内容的图像检索,能通过调整各个分块的权值来进一步提高检索的精度。

文献[70-71]采用固定尺寸的网格进行图像划分,提取分块的局部颜色直方图特征作为颜色空间特征,然而随着图像分块的数量增加,特征维数非常高。

2.3.7 基于投影的颜色空间特征

牛蕾等^[72]从 CT 图像重建的理论中得到启发,提出了基于 Radon 投影变换的图像检索技术。该方法基本思想是对一幅图像的 RGB 三个分量分别进行 Radon 投影变换,对图像在一个方向上做投影后,投影包含着图像在这个方向上的颜色密度信息;对图像在多个方向上做投影,投影向量组就有了一定的颜色空间信息。但是, Radon 变换是很耗时的,由于图像检索对实时性要求很高,这种图像检索的方法不够满意。因此,文献[72]提出分别将 RGB 三个分量进行小波变换单层分解,针对小波分解的低频信号进行 Radon 投影变换,进一步利用欧氏距离计算投影矩阵的相似性度量。

$f(x,y)$ 函数的 Radon 变换可以定义为沿包含该函数的平面的一族



直线的线积分，即

$$R\{f(x, y)\} = \iint f(x, y) \delta(t - x \cos \theta - y \sin \theta) dx dy = g_\theta(t) \quad (2-37)$$

式中， t 为沿直线上的距离，表示为 $t = x \cos \theta + y \sin \theta$ 沿着一系列平行线（投影）的积分就组成了投影 $g_\theta(t)$ ，所有的投影组成的集合 $\{g_\theta(t), \theta \in [0, \pi]\}$ 就是 Radon 变换。图 2-5 所示为 Radon 变换示意图。

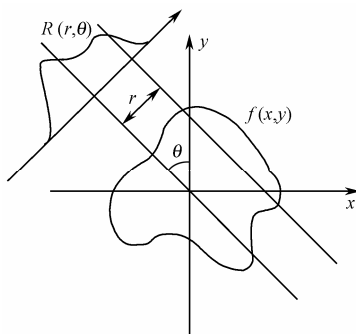


图 2-5 Radon 变换示意图

然而文献[72]的颜色空间特征不具有平移、旋转和尺度不变性，张志安等^[73]提出了改进算法，利用图像投影的质心对投影矩阵进行中心化以消除平移影响。进一步求 Radon 投影变换矩阵的奇异值分解，而矩阵的奇异值分解本生具有旋转不变性，这样 Radon 投影变换的颜色空间特征就具有平移、旋转和尺度不变的特性。

2.3.8 环形颜色空间特征

为了提取图像的颜色空间特征，Rao^[74]等人提出了一种环形颜色空间划分方法。设 $C_i = (x_i, y_i)$ 为图像中颜色为 i 的所有像素的质心， $|H_i|$ 表示集合 H_i 中像素的数目， x_i 和 y_i 分别定义为

$$x_i = \frac{1}{|H_i|} \sum_{(x,y) \in H_i} x_i \cdot y_i = \frac{1}{|H_i|} \sum_{(x,y) \in H_i} y_i \quad (2-38)$$

设 r_i 表示图像中颜色为 i 的像素同其质心的最大距离，即

$$r_i = \max \left[(x - x_i)^2 + (y - y_i)^2 \right]^{1/2} \quad (2-39)$$



可以设定的一个正整数 N ，把 r_i 分为 N 等份，然后以质心 C_i 为圆心，以 $(j \times r_i) / N$ 为半径 ($1 \leq j \leq N$) 画圆得到 N 个环。由内至外每个环与 H_i 的交点把 H_i 分为 $H_{i1}, H_{i2}, \dots, H_{iN}$ 。把 $H_{i1}, H_{i2}, \dots, H_{iN}$ 称为 H 的一个分割，该环形分割具有平移和旋转不变性。图 2-6 所示为等距环形颜色分割示意图。

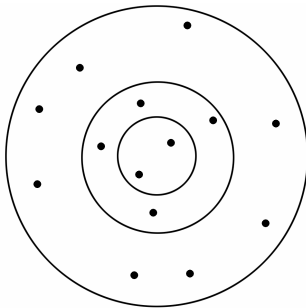


图 2-6 等距环形颜色分割示意图

由该方法可以统计图像的环形颜色直方图，可以看出环形颜色直方图具有平移和旋转不变性，进一步就可以进行特征向量的相似性度量。实验结果表明，环形颜色直方图比传统的颜色直方图具有较好的检索效果。

环形颜色直方图随着颜色分割的空间数量，特征维数非常高，这将大大增加存储空间的检索速度，因此孙君顶等^[58]提出环形颜色空间分布熵，颜色 i 的空间分布熵表示为

$$E_i = - \sum_{j=1}^n p_{ij} \log_2(p_{ij}) \quad (2-40)$$

式中， p_{ij} 为第 i 个颜色第 j 个环形分割空间的概率分布密度，即归一化后的颜色直方图，这样图像的特征可表示为 $F = \{(H_1, E_1), \dots, (H_i, E_i), \dots, (H_n, E_n)\}$ ，其中 H_i 为颜色直方图特征，而 E_i 为颜色信息熵特征。



2.3.9 广义颜色空间分布矩

视觉感知系统发现人类往往对颜色块更感兴趣, 环形颜色空间特征随着环形颜色分割数量的增加, 会导致空间颜色直方图特征向量变得较为庞大, 在一定程度上会导致检索性能的下降。因此, 文献[75]进一步提出了基于广义图像的空间颜色分布矩特征, 设彩色图像为 $f(x, y)$, 将 $f(x, y)$ 图像经过 3×3 点阵平滑后的图像为 $g(x, y)$, 由 $f(x, y)$ 和 $g(x, y)$ 构成一个二元组 (f, g) , 即为图像 $f(x, y)$ 的广义图像。图像的广义空间颜色分布矩在一定程度上表征了颜色块的统计特征, 符合人的视觉特征。首先分析原始图像 $f(x, y)$, 设 $h_i = (h_{i1}, h_{i2}, \dots, h_{iN})$ 为颜色 i 的空间分布直方图, N 表示第 N 个环形颜色分割, 由于图像的中心矩具有平移不变性, 为了降低计算复杂度, 下面采用三阶中心矩来计算:

$$\begin{cases} M_{i1}^f = \frac{1}{N} \sum_{j=1}^N h_{ij} \\ M_{i2}^f = \left[\frac{1}{N} \sum_{j=1}^N (h_{ij} - M_{i1}^f)^2 \right]^{1/2} \\ M_{i3}^f = \left[\frac{1}{N} \sum_{j=1}^N (h_{ij} - M_{i1}^f)^3 \right]^{1/3} \end{cases} \quad (2-41)$$

由环形颜色分割的平移和旋转不变性可知, M_{i1}^f , M_{i2}^f 和 M_{i3}^f 具有平移和旋转不变性, 为了获得尺度不变性, 可以将直方图进行归一化, 即 $h_{ij} = |H_{ij}| / |H_i|$ 表示为颜色 i 在第 j 个环形分割区域的归一化颜色直方图, 这样可以使直方图具有尺度不变性, 计算上述的矩特征即具有尺度不变性。

由于特征向量内部各分量的含义不同, 需要将不同的特征向量进行内部归一化, 使特征向量的内部各分量在相似性度量时地位相同, 可以采用高斯归一化^[51]的方法进行处理。若广义图像的 $f(x, y)$ 分量特征序列为 $F = [M_{1,j}^f, M_{i,j}^f, \dots, M_{w,j}^f]$, $M_{i,j}^f$ 表示第 i 幅图像第 j 个矩特征向量, w 表示图像库的图像数量。利用其均值 m_j 和标准差 σ_j , 可以将原序列



归一化为 $N(0,1)$ 分布的序列, 即

$$M_{i,j}^f = \frac{M_{i,j}^f - m_j}{\sigma_j} \quad (2-42)$$

同理, 可以针对广义图像 $g(x,y)$ 分量进行归一化。假设待检索的图像为 Q , P 为图像库中任意一目标图像, 针对广义图像的 $f(x,y)$ 分量环形颜色分布矩特征, 下面采用 L_1 -norm 距离来进行特征向量间的相似性度量。

$$d_f(Q,P) = \sum_{i=1}^n \left[\left| M_{i1}^{f,Q} - M_{i1}^{f,P} \right| + \left| M_{i2}^{f,Q} - M_{i2}^{f,P} \right| + \left| M_{i3}^{f,Q} - M_{i3}^{f,P} \right| \right] \quad (2-43)$$

同理, 可以针对广义图像 $g(x,y)$ 分量环形颜色分布矩特征, 仍然采用 L_1 -norm 距离来进行特征向量间的相似性度量, 即

$$d_g(Q,P) = \sum_{i=1}^n \left[\left| M_{i1}^{g,Q} - M_{i1}^{g,P} \right| + \left| M_{i2}^{g,Q} - M_{i2}^{g,P} \right| + \left| M_{i3}^{g,Q} - M_{i3}^{g,P} \right| \right] \quad (2-44)$$

对于所提取的不同图像特征, 需要将不同子特征计算出的距离进行线性组合, 由于不同的子特征采用的度量方式不同。因此, 为了使不同的子特征距离具有可比性, 需要对距离进行外部归一化, 算法仍然采用进行高斯归一化的方法。外部归一化的实质是对图像库中所有图像之间的相似距离进行归一化, 设检索图像库中所有图像间的距离所形成数列为 $D = [d_{1,2}, d_{1,3}, \dots, d_{i,j}, \dots, d_{w,w-1}]$, 其中 $d_{i,j}$ 为第 i 幅图像和第 j 幅图像的相似度, 序列的均值为 m_d 和标准差为 σ_d , 则外部归一化后的距离为

$$d_{i,j}^n = \left(\frac{d_{i,j} - m_d}{3\sigma_d} + 1 \right) / 2 \quad (2-45)$$

因此, 将广义图像 $f(x,y)$ 分量的距离 d_f 和广义图像 $g(x,y)$ 分量的距离 d_g 进行外部归一化后的距离分别为 d_1^n 和 d_2^n , 进一步两幅图像间的相似度量距离可定义为

$$D(P,Q) = \omega_1 d_1^n + \omega_2 d_2^n \quad (2-46)$$

其中 ω_1 , ω_2 为子特征的权重, 用于控制广义图像 $f(x,y)$ 和 $g(x,y)$ 分量在检索中的相对重要性, 并且满足 $\omega_1 + \omega_2 = 1, \omega_1, \omega_2 \in [0,1]$ 。



广义环形颜色空间分布矩检索算法如下所述。

【Step1】计算检索图像的广义图像模型，即 $f(x,y)$ 图像经过 3×3 点阵平滑后的图像为 $g(x,y)$ ，由 $f(x,y)$ 和 $g(x,y)$ 构成一个二元组 (f,g) 。

【Step2】针对广义图像的两个分量分别进行环形分割，计算颜色空间分布矩特征，以此来描述图像的颜色空间信息。

【Step3】对特征向量内部进行高斯归一化。

【Step4】计算广义图像分量 $f(x,y)$ 的颜色空间分布矩特征向量间的距离，本章采用 L_1 -norm 距离进行相似性度量，同理计算广义图像分量 $g(x,y)$ 的特征向量间的距离。

【Step5】将特征向量的距离进行外部归一化，按照式 (2-46) 计算图像间的综合距离作为相似度，并进行图像检索。

为了评价算法的性能，从检索图像库 (<http://wang.ist.psu.edu/>) 选取了 2000 幅图像作为检索测试库，其中包括动物、建筑、花卉和球场等图像，其中各类图像均超过 30 幅。实验中，先在图像库中选取 8 类图像，且在每类图像中随机抽取 10 幅图像共组成 80 次查询，然后结合这 80 次查询，计算出在不同检索率条件下系统的平均精确度。算法采用 3×3 点阵平滑得到广义图像，然后将环形颜色分割设为 5 个环形分割区域，将权重设置为 $\omega_1 = \omega_2 = 0.5$ ，即广义图像的两个分量权重设置为相等的地位。

所有实验在相同的软硬件环境下实现，将广义颜色空间分布矩算法 (Generalized patial-color moment algorithm, GPCM)、Geostat 算法、SCH 算法进行了检索性能比较，三种算法的查全率与查准率变化曲线如图 2-7 所示。由图 2-7 可以看出，广义颜色空间分布矩 GPCM 算法的检索性能明显优于其他两种算法。由于 Geostat 方法和 SCH 算法反映的是孤立的像素点在图像空间中的分布特征，而广义颜色空间分布矩 GPCM 算法提取了图像颜色块的空间统计特征，进一步细化了图像的颜色空间



信息，因此检索效果较好。广义颜色空间分布矩 GPCM 算法利用每一种颜色的颜色空间矩表征图像的颜色空间特征，并且利用了广义图像的两个分量进行提取特征，因此检索效果在一定程度上有所提高。图 2-8 所示为广义环形颜色空间分布矩检索算法针对公交车的检索示例图。

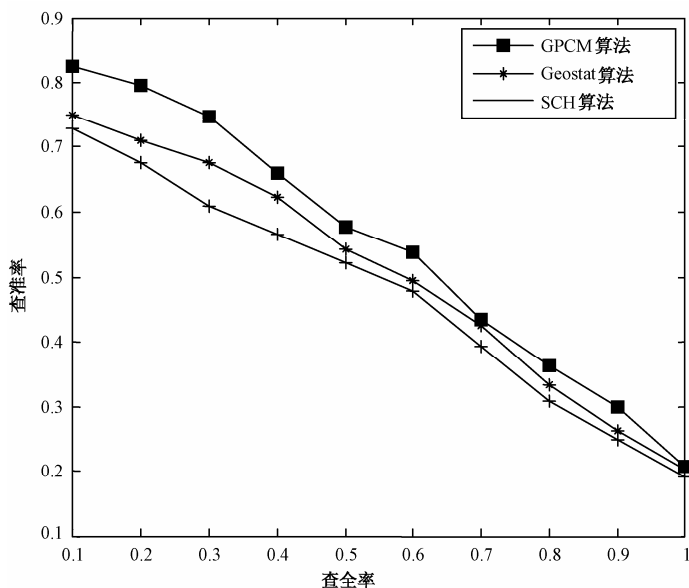


图 2-7 三种算法的检索性能



图 2-8 广义环形颜色空间分布矩检索示例图



2.4 本章小结

本章首先详细介绍了颜色空间模型,然后介绍了经典基于颜色特征的图像检索算法,其中包括颜色直方图、颜色统计矩、颜色信息熵和颜色视觉关注度等相关算法,最后介绍了基于颜色和空间特征的图像检索算法。由于传统的颜色方法描述了图像的全局颜色特征,不包括图像颜色的空间分布特征,因此仅利用这些特征进行图像检索时极易造成误检现象,因此较多基于颜色和空间特征的图像检索算法相继被提出,其中包括广义颜色直方图、颜色集、颜色聚合向量、颜色相关图、SCH 算法、geostat 算法、局部空间特征算法、环形颜色空间特征和颜色空间分布矩算法等。实验分析表明,基于颜色和空间特征的图像检索算法的检索性能要优于基于颜色特征的检索算法。

第三章



基于形状特征的图像检索

3.1 形状特征概述

形状是图像的一种基本视觉特征，形状信息不随图像颜色的变化而变化，是物体稳定的特征之一。研究表明，人的视觉系统对图像中的物体的形状更敏感，人们只凭物体的形状信息就可以识别物体，这也是形状有别于颜色、纹理等视觉特征的关键所在。因此，准确地提取图像中物体的形状特征，并将其与其他图像特征一起有效地应用于图像检索中具有重要意义，必将大大提高图像检索的准确率，基于形状的图像检索研究成为 CBIR 中一个极具挑战性的研究课题。

图像的形状描述必须满足以下特点^[76]。

- 1) 独特性 指每幅图像必须具有一个独特的描述；
- 2) 完备性 指描述尽量准确完整；
- 3) 灵敏性 容易反映相似目标的差异；
- 4) 几何变换不变性 当图像发生平移、旋转、缩放和镜像变换后，形状描述具有不变性；
- 5) 抽象性 指形状描述要能从细节中抽象出形状的基本特征。

目前，形状特征的描述主要分为基于区域的方法和基于边界的方法两大类。基于区域的方法是利用形状的全部区域进行研究，而基于边界



的方法则是利用形状的外部边缘进行研究。本章将介绍这两类形状分析方法。

3.2 基于区域的形状特征分析

基于区域的形状特征分析利用整个区域的所有像素进行分析, 具有较好的检索性能。常见的方法包括区域形状参数、HU 不变矩、Legendre 矩、Tchebichef 矩、Zernike 矩和 Krawtchouk 矩等。

3.2.1 区域形状参数

常见的图像区域描述方法主要包括区域面积、区域重心、区域拓扑参数、形状参数、偏心率、球状性、圆形性和实体性等^[76]。

1) 区域面积 计算区域面积就是对属于区域的像素计数。

$$A = N \quad (3-1)$$

2) 区域重心 区域重心的坐标是根据所有属于区域的点计算出来的。

$$\bar{x} = \frac{1}{A} \sum_{(x,y) \in R} x \quad (3-2)$$

$$\bar{y} = \frac{1}{A} \sum_{(x,y) \in R} y \quad (3-3)$$

式中, A 为区域面积。

3) 拓扑描述符

(1) 拓扑学 (Topology): 是研究图形不受畸变变形 (不包括撕裂或粘贴) 影响的性质的。

(2) 欧拉数:

$$E = C - H \quad (3-4)$$

式中, C 为连通组元的个数, H 为孔数。

(3) 多边形网的欧拉公式:



$$W-Q+F=E=C-H \quad (3-5)$$

式中, W 为顶点数, Q 为边数, F 为面数。

4) 形状描述符

(1) 形状参数:

$$F = \frac{\|B\|^2}{4\pi A} \quad (3-6)$$

式中, B 为区域的周长, A 为区域的面积。此参数在一定程度上描述了区域的紧凑性, 对尺度和旋转不敏感。

(2) 偏心率: 也称为伸长度, 有如下两种计算方法: 计算边界长轴(直径)长度和短轴长度的比值, 此结果受物体形状和噪声的影响比较大; 另一种方法是用整个区域的所有像素, 抗噪声等干扰的能力较强。

(3) 球状性:

$$S = \frac{r_i}{r_c} \quad (3-7)$$

式中, r_i 为区域的内切圆的半径, r_c 为区域的外切圆的半径。

(4) 圆形性:

$$C = \frac{M_R}{\delta_R} \quad (3-8)$$

$$M_R = \frac{1}{N} \sum_{k=0}^{N-1} \|(x_k - y_k) - (\bar{x} - \bar{y})\| \quad (3-9)$$

$$\delta_R = \frac{1}{N} \sum_{k=0}^{N-1} [\|(x_k - y_k) - (\bar{x} - \bar{y})\| - M_R]^2 \quad (3-10)$$

式中, $(x_k - y_k)$ 为区域边界点, $(\bar{x} - \bar{y})$ 为重心, 当区域为圆形时, C 单增趋向无穷。圆形性不受区域平移、旋转和尺度变换的影响。

(5) 实体性:

$$\text{Solidity} = \frac{A}{H} \quad (3-11)$$

式中, A 为多边形的面积, H 为多边形的凸壳面积, Solidity 描述了形状的凹凸性。



3.2.2 HU不变矩

图像的 HU 不变矩特征是描述区域形状较为有效的方法。假设 R 是用二值图像表示的物体，则 R 形状的第 $p+q$ 阶中心矩为

$$\mu_{p,q} = \sum_{(x,y) \in R} (x-x_c)^p (y-y_c)^q f(x,y) \quad (3-12)$$

式中， (x_c, y_c) 是物体的中心。为获得与缩放无关的性质，可以对该中心矩进行标准化操作：

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^\gamma}, \gamma = \frac{p+q+2}{2} \quad (3-13)$$

基于这些矩，HU^[77]提出了以下 7 个不变矩：

$$\phi_1 = \log(\mu_{2,0} + \mu_{0,2}) \quad (3-14)$$

$$\phi_2 = \log((\mu_{2,0} - \mu_{0,2})^2 + 4\mu_{1,1}^2) \quad (3-15)$$

$$\phi_3 = \log((\mu_{3,0} - 3\mu_{1,2})^2 + (\mu_{0,3} - 3\mu_{2,1})^2) \quad (3-16)$$

$$\phi_4 = \log((\mu_{3,0} + \mu_{1,2})^2 + (\mu_{0,3} + \mu_{2,1})^2) \quad (3-17)$$

$$\phi_5 = \log\left((\mu_{3,0} - 3\mu_{1,2})(\mu_{3,0} + \mu_{1,2})\left[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{0,3} + \mu_{2,1})^2\right] + (\mu_{0,3} - 3\mu_{2,1})(\mu_{0,3} + \mu_{2,1})\left[(\mu_{0,3} + \mu_{2,1})^2 - 3(\mu_{3,0} + \mu_{1,2})^2\right]\right) \quad (3-18)$$

$$\phi_6 = \log\left((\mu_{2,0} - \mu_{0,2})\left[(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{0,3} + \mu_{2,1})^2\right] + 4\mu_{1,1}(\mu_{3,0} + \mu_{1,2})(\mu_{0,3} + \mu_{2,1})\right) \quad (3-19)$$

$$\phi_7 = \log\left((3\mu_{2,1} - \mu_{0,3})(\mu_{3,0} + \mu_{1,2})\left[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{0,3} + \mu_{2,1})^2\right] + (\mu_{0,3} - 3\mu_{2,1})(\mu_{0,3} + \mu_{2,1})\left[(\mu_{0,3} + \mu_{2,1})^2 - 3(\mu_{3,0} + \mu_{1,2})^2\right]\right) \quad (3-20)$$

HU 利用二阶和三阶规格化中心矩导出了 7 个具有平移、缩放和旋转不变性的不变矩，然而图像的二阶和三阶中心矩对噪声等因素非常敏感，因此 HU 不变矩不适合于图像噪声比较大的情况，同时 HU 不变矩存在一定的冗余信息。

图 3-1 所示为飞机形状示意图，根据 HU 不变矩算法可以得到 7 个矩特征，表 3-1 为图 3-1 的 HU 不变矩特征向量。



图 3-1 飞机形状示意图

表 3-1 图 3-1 的 HU 不变矩特征向量

ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7
-9.6031	-20.2888	-32.6664	-32.7016	-65.3245	-42.9756	-65.7799

3.2.3 Legendre矩

为了克服了 HU 不变矩存在冗余信息的缺点, Teague^[78]提出了正交 Legendre 矩理论。对于一幅密度函数为 $f(x, y)$ 的二维图像, 其 $p+q$ 阶的 Legendre 正交矩的定义为

$$L_{pq} = \frac{(2p+1)(2q+1)}{4} \int_{-1}^1 \int_{-1}^1 P_p(x) P_q(y) f(x, y) dx dy \quad (3-21)$$

式中, $P_p(x)$ 为 p 阶的 Legendre 多项式, 其定义为

$$P_p(x) = \frac{1}{2^p} \sum_{k=0}^{n/2} (-1)^k \frac{(2p-2k)}{k!(p-k)!(p-2k)!} x^{p-2k}, x \in [-1, 1] \quad (3-22)$$

对于一幅 $N \times N$ 的离散化数字图像, 式 (3-16) 可用下式逼近,

$$L_{pq} = \frac{(2p+1)(2q+1)}{(N-1)^2} \sum_{i=1}^N \sum_{j=1}^N P_p(x_i) P_q(y_j) f(x_i, y_j) \quad (3-23)$$

式中, $x_i = (2i - N - 1)/(N - 1)$, $y_i = (2j - N - 1)/(N - 1)$ 。Legendre 多项式满足如下迭代性质:

$$P_{p+1}(x) = \frac{2p+1}{p+1} x P_p(x) - \frac{p}{p+1} P_{p-1}(x), p \geq 1 \quad (3-24)$$



式中, $P_0(x)=1, P_1(x)=x$ 。

由于 Legendre 矩的计算复杂度较高, 因此杨冠羽等^[79]提出了基于灰度图像 Legendre 矩的快速算法, 降低了计算复杂度。马丽红等^[80]成功将其应用于虹膜识别算法中, 具有较好的识别效果。

3.2.4 Zernike矩和Pseudo-Zernike矩

为了克服 HU 不变矩针对噪声敏感性等问题, 文献[81]提出了 Zernike 矩。实验证明, Zernike 矩在信息冗余度和识别率等均优于 HU 不变矩和 Legendre 矩, 已被广泛应用于模式识别领域。Zernike 提出了一组定义在极坐标下 ($x^2 + y^2 \leq 1$) 下的复数多项式 $V_{nm}(x, y)$,

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = \exp(jm\theta) \quad (3-25)$$

式中, n 为正整数或零, m 为满足条件, $n - |m|$ 为偶数, 且 $|n| \geq m$, (ρ, θ) 是点 (x, y) 的极坐标表示, ρ 表示从圆心到点 (x, y) 的向量长度, θ 为向量 ρ 沿逆时针方向与 x 轴的夹角, $f(x, y)$ 为向量长度, $f(x, y)$ 等于 0 或 1。 $R_{nm}(\rho)$ 为正交的径向多项式, 且有 $R_{n,m}(\rho) = R_{n,-m}(\rho)$, 定义为

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} (-1)^s \cdot \frac{(n-s)!}{s! \left(\frac{n-|m|}{2} - s \right)! \left(\frac{n+|m|}{2} - s \right)!} \rho^{n-2s} = \sum_{k=1}^n B_{nmk} \rho^k \quad (3-26)$$

在 Zernike 正交多项式基础上的 n 阶 Zernike 定义为

$$Z_{nm} = \left[\frac{n+1}{\pi} \right] \iint_{x^2+y^2 \leq 1} V_{nm}^*(x, y) f(x, y) dx dy \quad (3-27)$$

式中, $V_{nm}^*(x, y)$ 为 $V_{nm}(x, y)$ 的共轭, 对于一幅离散数字图像 $f(x, y)$ 的 n 阶 Zernike 矩表示为

$$Z_{nm} = \left[\frac{n+1}{\pi} \right] \sum_x \sum_y V_{nm}^*(x, y) f(x, y), x^2 + y^2 \leq 1 \quad (3-28)$$

Zernike 矩具有很好的旋转不变性, 已被广泛应用于图像处理和模式识别领域^[82-84]。为了得到具有平移和尺度不变的形状特征, 在计算



Zernike 矩前要对图像进行归一化操作,同时要将图像的重心移到图像的中心,平移和尺度不变性可通过下式获得:

$$g(x, y) = f\left(\frac{x}{a} + \bar{x}, \frac{y}{a} + \bar{y}\right) \quad (3-29)$$

式中, (\bar{x}, \bar{y}) 为图像函数 $f(x, y)$ 的重心, $a = \sqrt{\beta / m_{00}}$, β 为预先定义的值, $\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$, m_{00} 、 m_{10} 和 m_{01} 分别为图像函数的几何矩。 m_{00} 、 m_{10} 和 m_{01} 矩的定义为

$$m_{00} = \iint f(x, y) dx dy \quad (3-30)$$

$$m_{01} = \iint y f(x, y) dx dy \quad (3-31)$$

$$m_{10} = \iint x f(x, y) dx dy \quad (3-32)$$

Teh 和 Chin^[85]将 Zernike 正交多项式稍作修改,得到 Pseudo-Zernike 多项式:

$$R_{nm}(r) = \sum_{s=0}^{\frac{n-|m|}{2}} (-1)^s \cdot \frac{(2n+1-s)!}{s!(n-|m|-s)!(n+|m|+1-s)!} r^{n-s} \quad (3-33)$$

式中, n 为正整数或零, m 为满足条件, $n-|m|$ 为偶数, 且 $|n| \geq m$, 则阶数为 n , 重复度为 m 的 Pseudo-Zernike 矩定义为

$$A_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 R_{nm}(r) \cdot \exp(-j l \theta) f(r, \theta) r dr d\theta \quad (3-34)$$

Pseudo-Zernike 矩描述子具有以下优点^[86-88]: ①旋转不变性; ②高斯噪声鲁棒性; ③冗余性小; ④能描述运动图像, 不仅可用于静止图像的特征提取, 而且还可以用于运动图像分析; ⑤易于任意构造高阶矩, 而不像 Hu 矩要去找具有不变性的高阶矩那样非常困难。由此可见, Pseudo-Zernike 矩是一个理想的形状描述算子, 然而 Pseudo-Zernike 矩描述子不具有尺度不变性, 因此在进行特征提取前可以将图像进行归一化处理, 这样 Pseudo-Zernike 矩就具有良好的识别能力。



3.2.5 Tchebichef 矩

为了提高图像识别能力, Mukundan^[89]提出了基于 Tchebichef 离散正交多项式的 Tchebichef 矩。由于 Tchebichef 正交矩本身是离散的, 不需要对定义域进行归一处理, 可直接应用于离散图像, 计算精度较高; Tchebichef 多项式的计算具有递推关系和对称性, 因此 Tchebichef 矩能够进行快速计算和图像重建。设大小为 $N \times N$ 的原始图像 $f(x, y)$, 则离散 Tchebichef 多项式为^[89,90]

$$t_n(x) = (1-N)_n \times {}_3F_2(-n, -x, 1-N; 1) \quad (3-35)$$

式中, $n, x = 0, 1, \dots, N-1$, ${}_2F_3(a_1, a_2, a_3; b_1, b_2; z) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k (a_3)_k z^k}{(b_1)_k (b_2)_k k!}$, $(a)_k = a(a+1)(a+2)\dots(a+k-1)$, 以 Tchebichef 多项式的乘积作为核函数引出 Tchebichef 矩为

$$T_{pq} = \frac{1}{\rho(p, N)\rho(q, N)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} t_p(x) t_q(y) f(x, y) \quad (3-36)$$

式中, $p, q = 0, 1, \dots, N-1$, $\rho(p, N)$ 的定义为

$$\rho(p, N) = \frac{N^2(N^2-1)(N^2-2)\dots(N^2-n^2)}{2n+1} \quad (3-37)$$

式中, $p, q = 0, 1, \dots, N-1$ 。从上面 Tchebichef 矩的定义可知, 其定义域与离散图像的取值范围相吻合。但随着阶数的增大, Tchebichef 矩的取值范围过大, 不宜进行集中分析, 因此要对其值进行加权修正。加权的 Tchebichef 多项式为

$$\tilde{t}_n(x) = \frac{t_n(x)}{\beta(n, N)} = \sum_{k=0}^n c_{k, n, N} x^k \quad (3-38)$$

式中, $c_{k, n, N} = (-1)^{n-k} \binom{N-1-K}{n-k} \binom{n+k}{n}$, $\beta(n, N) = N^n$ 。这样加权的

Tchebichef 矩的表达式为

$$T_{pq} = \frac{1}{\tilde{\rho}(p, N)\tilde{\rho}(q, N)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \left(\sum_{i=0}^p c_{i, p, N} x^i \right) \left(\sum_{j=0}^q c_{j, q, N} x^j \right) f(x, y) \quad (3-39)$$



式中, $\tilde{\rho}(n, N) = \frac{\rho(n, N)}{\beta(n, N)^2} = \frac{N \left(1 - \frac{1}{N^2}\right) \left(1 - \frac{2^2}{N^2}\right) \cdots \left(1 - \frac{n^2}{N^2}\right)}{2n+1}$ 。上述过程也可

利用几何矩来进行计算, 为了获得平移、缩放、旋转均不变的 Tchebichef 矩, 可分别通过对 $f(x, y)$ 进行平移、缩放和旋转处理得到。

实验表明, Tchebichef 矩的检索率优于 HU 不变矩、Zernike 矩和 Legendre 矩的性能^[91]。

3.2.6 Krawtchouk 矩

尽管 Legendre 矩、Zernike 矩和 Tchebichef 矩在检索算法中具有较好的检索性能, 但这些方法仍然在数字图像领域存在离散化、坐标转换及构造复杂等问题。为了克服这些问题, Yap^[92]等人提出了 Krawtchouk 矩, 实验结果表明, 与 Zernike 矩相比, Legendre 矩和 Tchebichef 矩具有较好的识别效果。Krawtchouk 矩的正交变换核为 Krawtchouk 多项式, n 阶 Krawtchouk 多项式的定义为

$$K_n(x; p, N) = \sum_{k=0}^n a_{k,n,p} x^k = {}_2F_1 \left(-n, -x; -N; \frac{1}{p} \right), x, n = 0, 1, 2; N > 0; p \in (0, 1) \quad (3-40)$$

式中, ${}_2F_1$ 的表达式为

$${}_2F_1(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!} \quad (3-41)$$

$(a)_k$ 的表达式为

$$(a)_k = a(a+1) \cdots (a+k-1) = \frac{\Gamma(a+k)}{\Gamma(a)} \quad (3-42)$$

假设图像大小为 $M \times N$, 给定图像函数 $f(x, y)$, $m+n$ 阶 Krawtchouk 矩的定义为

$$K_{mn} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \overline{k_n(x; p_1, N-1)} \overline{k_m(y; p_2, M-1)} f(x, y) \quad (3-43)$$

式中,



$$\begin{cases} \overline{k_n}(x; p_1, N-1) = k_n(x; p_1, N-1) \sqrt{\frac{\omega(x; p, N-1)}{\rho(n; p, N-1)}} \\ \omega(x; p, N) = \binom{N}{x} p^x (1-p)^{N-x} \\ \rho(n; p, N) = \frac{(-1)^n}{(-N)_n} \left(\frac{1-p}{p} \right)^n \end{cases} \quad (3-44)$$

称 $\overline{k_n}()$ 为加权 Krawtchouk 多项式, Krawtchouk 矩的定义域为 $[(0,0), (M-1, N-1)]$, 正好是图像所在的空间坐标系, 因此进行计算时不需要进行坐标变换, 适合用于描述数字域的信号。为了获得 Krawtchouk 矩的几何不变性, Yap 等人利用 Krawtchouk 矩和形状不变矩的关系得到了具有平移、旋转和尺度不变的 Krawtchouk 不变矩。

3.3 基于边界轮廓的形状分析

基于边界轮廓的形状分析是在图像边缘或轮廓检测基础上, 进行形状特征提取, 因此边缘检测技术在基于边界轮廓的形状分析中具有重要的地位。

3.3.1 傅里叶形状描述符

傅里叶变换能将满足一定条件的某个函数表示成三角函数(正弦和/或余弦函数)或其积分的线性组合。傅里叶变换广泛应用于图像变换、图像编码与压缩、图像分割等领域中。傅里叶描述符(Fourier Shape Descriptors)^[93]的基本思想是用物体边界的傅里叶变换作为其形状描述, 设二维物体的轮廓由一系列坐标为 (x_s, y_s) 的像素组成, 其中 $0 \leq s \leq N-1$, N 为轮廓上像素的总数, 可以得到形状边界的曲率函数(Curvature Function)、质心距离(Centroid Distance)和复坐标函数(Complex Coordinates Function)等形状特征。



1. 曲率函数

曲率函数为图像轮廓切向角度相对于弧长的变化率。曲率函数 $K(s)$ 可以表示为

$$K(s) = \frac{d}{ds} \theta(s) \quad (3-45)$$

式中, $\theta(s)$ 为轮廓线的切向角度,

$$\left. \begin{aligned} \theta(s) &= \arg \tan \left(\frac{y'_s}{x'_s} \right) \\ y'_s &= \frac{dy_s}{ds} \\ x'_s &= \frac{dx_s}{ds} \end{aligned} \right\} \quad (3-46)$$

针对曲率函数和质心距离函数而言, 函数的傅里叶变换是对称的, 可以只考虑正频率的坐标轴, 基于曲率函数的形状描述符表示为

$$f_K = [|F_1|, |F_2|, \dots, |F_{M/2}|] \quad (3-47)$$

式中, F_i 为傅里叶变换参数的第 i 个分量。

2. 质心距离

质心距离为从物体边界点到物体中心 (x_c, y_c) 的距离, 即

$$R(s) = \sqrt{(x_s - x_c)^2 + (y_s - y_c)^2} \quad (3-48)$$

质心距离的傅里叶形状描述符为

$$f_R = \left[\frac{|F_1|}{|F_0|}, \frac{F_2}{F_0}, \dots, \frac{|F_{M/2}|}{|F_0|} \right] \quad (3-49)$$

式中, F_i 为傅里叶变换参数的第 i 个分量, 式 (3-48) 是欧氏距离, 也可以定义其他的距离测度。

3. 复坐标函数

复坐标函数是用复数所表示的像素坐标, 即

$$Z(s) = (x_s - x_c) + j(y_s - y_c) \quad (3-50)$$



复坐标函数的傅里叶变换会产生一系列复数系数,系数的低频分量表征形状的宏观属性,高频分量表征形状的细节特征,形状描述符可以从这些变换参数中得出。为了获得尺度不变性和旋转不变性,可以将参数的大小除以 DC 分量,同时只保留参数的大小信息,而忽略相位信息。对于复坐标函数,正频率分量和负频率分量被同时采用。由于 DC 参数与形状的所处的位置有关而被省略。因此,第一个非零的频率分量被用于对其他变换参数进行标准化。复坐标函数所导出的形状描述符为

$$f_z = \left[\frac{|F_{-(M/2-1)}|}{|F_1|}, \dots, \frac{|F_{-1}|}{|F_1|}, \frac{|F_2|}{|F_1|}, \dots, \frac{|F_{M/2}|}{|F_1|} \right] \quad (3-51)$$

式中, F_i 表示傅里叶变换参数的第 i 个分量, $i = -(M/2-1), \dots, -1, 2, \dots, M/2$ 。为保证数据库中所有物体的形状特征都有相同的长度,在实施傅里叶变换前,需要将所有边界点的数目统一为 M 。

3.3.2 基于内角的形状特征

文献[8]提出了一种基于内角的形状特征表达方法,将物体近似地表达成多边形的形式,可以表达为 $\text{Intra-angle} = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 。内角的形状描述与形状所在位置、旋转和大小无关,因此内角的形状特征非常适合图像检索技术。下面是一系列从内角导出的形状特征描述符。

1. 顶点数

多边形的顶点数目越多,形状就越复杂。把具有不同顶点数目的两个形状当做很不相似的两个形状有一定的合理性。

2. 内角平均值

多边形所有内角的平均值从一定程度上反映了多边形的形状属性,如三角形的内角平均值为 30° ,与矩形的内角平均值 90° 之间有较大差别。



3. 内角标准方差

多边形内角的标准方差为

$$\delta = \sqrt{\sum_{i=1}^n (a_i - \bar{a})^2} \quad (3-52)$$

式中, \bar{a} 是内角的平均值。该标准方差 δ 是多边形的总体描述。多边形越规则, δ 值越小。因此, 它可以用于分辨正多边形和不规则多边形。

4. 内角直方图

内角直方图反映了内角的总体分布, 该方法将 $0^\circ \sim 360^\circ$ 的角度范围等分成 k 个区间, 作为直方图的 k 个 bin, 然后统计每个角度区间中的内角数目, 如图 3-2 所示。

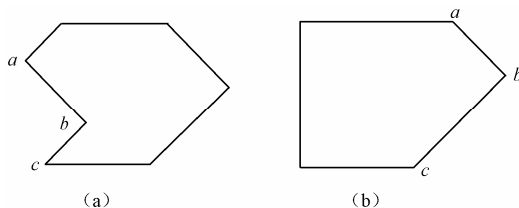


图 3-2 图像内角特征计算

下面介绍图 3-2 中内角 $\theta = \angle abc$ 的计算。设 a 、 b 、 c 三点的中心为 p , 则有

$$\vec{op} = \frac{\vec{oa} + \vec{ob} + \vec{oc}}{3} \quad (3-53)$$

式中, o 为原点。如 p 在多边形内部, 则 θ 小于 180° , 否则 θ 大于 180° 。

当 $\theta \leq 180^\circ$ 时,

$$\theta = \arccos \left(\frac{|ab|^2 + |bc|^2 - |ac|^2}{2|ab||bc|} \right) \quad (3-54)$$

当 $\theta > 180^\circ$ 时,

$$\theta = 360^\circ - \arccos \left(\frac{|ab|^2 + |bc|^2 - |ac|^2}{2|ab||bc|} \right) \quad (3-55)$$



3.3.3 小波HU不变矩

小波变换比傅里叶变换有更好的精度和稳定性,因此比较适合用于描述轮廓的基本特征。姚玉荣^[94]等提出采用小波模极大值来提取图像中的多尺度目标边缘信息,然后采用多尺度的不变矩来度量图像中目标形状的相似性等。对图像来说,小波模极大值描述的是图像中目标的多尺度边界^[95]。

小波模极大值是在对多尺度小波变换进行不规则抽样的基础上得到的,可以描述信号的奇异性,具有平移不变性。若 $g(x, y)$ 在整个平面上的积分为 1, 并且它在 x 或 y 为无限远处收敛到 0, 则定义 $g(x, y)$ 为二维平滑函数。进一步定义两个小波函数

$$w_1(x, y) = \frac{\partial g(x, y)}{\partial x}, \quad w_2(x, y) = \frac{\partial g(x, y)}{\partial y} \quad (3-56)$$

这样, 图像 $f(x, y)$ 小波变换的两个分量在尺度为 s 时定义

$$w_{1,s}f(x, y) = f * w_{1,s}(x, y) \quad w_{2,s}f(x, y) = f * w_{2,s}(x, y) \quad (3-57)$$

对于二进小波, 则

$$\begin{bmatrix} w_{1,2^j}f(x, y) \\ w_{2,2^j}f(x, y) \end{bmatrix} = 2^j \begin{bmatrix} \frac{\partial}{\partial x}(f * g_{2^j})(x, y) \\ \frac{\partial}{\partial y}(f * g_{2^j})(x, y) \end{bmatrix} = 2^j \nabla(f * g_{2^j})(x, y) \quad (3-58)$$

在任一尺度 2^j 上梯度的模为 $M_{2^j}f(x, y) = \sqrt{|w_{1,2^j}f(x, y)|^2 + |w_{2,2^j}f(x, y)|^2}$, 梯度与水平的夹角 $A_{2^j}f(x, y) = \arg[w_{1,2^j}f(x, y) + iw_{2,2^j}f(x, y)]$, $fg_{2^j}(x, y)$ 上剧烈的点是沿着梯度方向 $A_{2^j}f(x, y)$ 上模 $M_{2^j}f(x, y)$ 为局部极大值的那些点, 记录模极大值点的位置信息、 $A_{2^j}f(x, y)$ 和 $M_{2^j}f(x, y)$ 就可以描述图像的多尺度边界。对于小波分解层数的选择, Mallat 证明对于一幅 $N \times N$ 的图像, 最大分解层数 $J = \log_2(N) + 1$ 。实验表明, 小波模极大值在分解层数大于 6 时就几乎没有什么区分能力了。



HU 不变矩是基于区域的物体形状表示方法，而图像小波变换的模极大值点所对应的是边界（曲线），即边界的矩，区域矩的公式不能直接用于计算曲线矩，所以必须对不变矩的计算公式进行修正。对于区域 $f(x,y)$ 来说，若给它一个尺度 α , $x' = \alpha x$, $y' = \alpha y$ ，它的矩就要乘以 $\alpha^2 \alpha^p \alpha^q$ ，其中 α^2 是由于尺度变化而带来的目标面积变化所引起的， $f(x',y')$ 的中心矩就变成了 $u' = u_{pq} \times \alpha^{p+q+2}$ ，可以对该中心矩进行标准归一化操作

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^\gamma} \quad (3-59)$$

为了使得归一化矩对尺度变换是不变的，必须有 $\eta'_{pq} = \eta_{pq}$ ，即

$$\frac{u_{pq} \times \alpha^{p+q+2}}{(u_{00} \times \alpha^2)^\gamma} = \frac{u_{pq}}{(u_{00})^\gamma} \quad (3-60)$$

所以可推出

$$\gamma = p + q + 2 \quad (3-61)$$

对于曲线来说，尺度的变化造成周长的变化，变化因子为 α ，而不是 α^2 ，此时中心矩变成了 $u' = u_{pq} \times \alpha^{p+q+1}$ 。同样，为了使归一化矩具有尺度变换不变性，必须有 $\eta'_{pq} = \eta_{pq}$ ，即

$$\frac{u_{pq} \times \alpha^{p+q+1}}{(u_{00} \times \alpha)^\gamma} = \frac{u_{pq}}{(u_{00})^\gamma}, \text{ 所以可推出} \quad \gamma = p + q + 1 \quad (3-62)$$

可以看出式（3-61）与式（3-62）之间的不同，小波 HU 不变矩使用了修正后的不变矩的计算式（3-62）。

为了评价算法的检索性能，从 Corel 检索图像库选取了 1500 图像作为检索测试库，包括恐龙、花卉、马、大象、公交车各类图像，相似图像均超过 30 幅。图 3-3 所示为小波 HU 不变矩算法的检索结果。

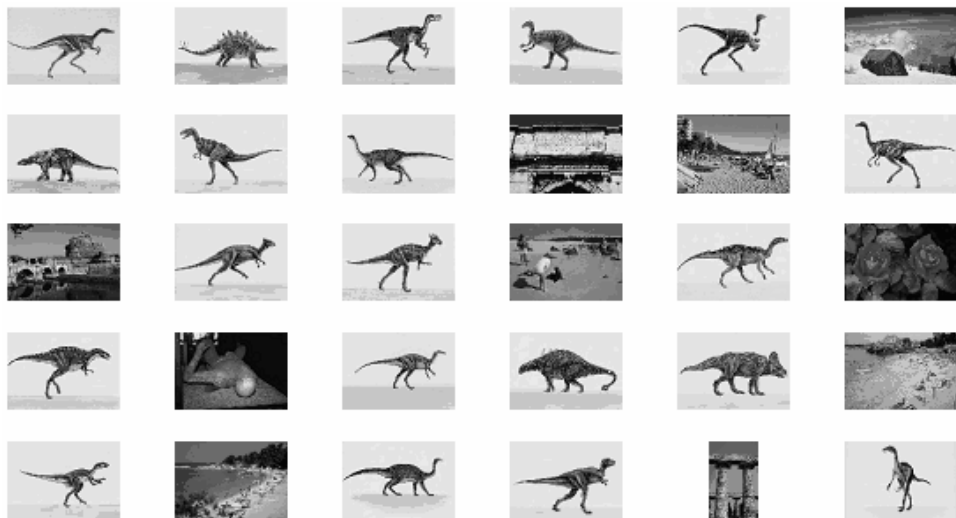


图 3-3 小波 HU 不变矩算法的检索结果

3.3.4 Radon轮廓不变矩

二维函数 $f(x, y)$ 的 Radon 变换是该函数沿包含该函数的平面的一族直线的线积分，即

$$R\{f(x, y)\} = \iint f(x, y) \delta(t - x \cos \theta - y \sin \theta) dx dy = g_{\theta}(t) \quad (3-63)$$

式中， t 代表沿直线上的距离，表示为 $t = x \cos \theta + y \sin \theta$ 沿着一系列平行线（投影）的积分就组成了投影 $g_{\theta}(t)$ ，所有的投影组成的集合 $\{g_{\theta}(t), \theta \in [0, \pi]\}$ 就是 Radon 变换。Radon 变换具有如下性质^[96]。

➤ 平移：设 x 和 y 方向的平移量为 t_x 和 t_y ，则

$$R(f(x - t_x, y - t_y)) = g_{\theta}(t - t_x \cos \theta - t_y \sin \theta) \quad (3-64)$$

可见原图像的平移就等于 t 的平移。

➤ 尺度：上标 s 表示尺度变换量， λ 表示比例因子。

$$R\{f(\lambda x, \lambda y)\} = g_{\theta}^s(t) = \frac{1}{\lambda} g_{\theta}(\lambda t) \quad (3-65)$$

➤ 旋转：设图像的极坐标表示为 $f(r, \theta)$ ， φ_0 为旋转量，则

$$R(f(r, \theta + \varphi_0)) = g(t, \theta + \varphi_0) \quad (3-66)$$



➤ 对称性:

$$g(t, \theta) = g(-t, \theta \pm \pi) \quad (3-67)$$

可见原图像的旋转就等于 θ 角度的平移。寻求具有平移、旋转和尺度不变的形状特征在检索系统中非常重要, 针对图像边缘计算其 Radon 几何不变特征向量, 因此也能够在一定程度上降低计算量。Li^[96]构造了基于区域的 Radon 不变矩, 而下面将针对边缘轮廓构造平移、旋转和尺度不变量。设图像 Radon 变换为 $g_\theta(t)$, 则定义 k 阶矩为

$$m_k(\theta) = \int t^k g_\theta(t) dt \quad (3-68)$$

进一步可以定义 $\bar{t}_\theta = \frac{m_1(\theta)}{m_0(\theta)}$ 表示 $g_\theta(t)$ 的质心。

1. 平移不变量

由于图像的平移会使投影产生平移, 用中心矩代替普通矩可以获得平移不变性。

$$\mu_k(\theta) = \int (t - \bar{t}_\theta)^k g_\theta(t) dt \quad (3-69)$$

2. 尺度不变量

由式 (3-69) 可知, $m_0^s(\theta) = m_0(\theta) / \lambda^2$, $m_1^s(\theta) = m_1(\theta) / \lambda^3$,

$$\bar{t}_\theta^s = \frac{m_1^s(\theta)}{m_0^s(\theta)} = \bar{t}_\theta / \lambda \quad (3-70)$$

则 $g_\theta^s(t)$ 的 k 阶中心矩为

$$\mu_k^s(\theta) = \int (t - \bar{t}_\theta^s)^k g_\theta(\lambda t) / \lambda dt = \int (\alpha - \bar{t}_\theta)^k g_\theta(\alpha) d(\alpha) / \lambda^{2+k} = \mu_k(\theta) / \lambda^{2+k} \quad (3-71)$$

当图像发生尺度变换时, 设 λ 为尺度比例因子, Radon 投影矩阵的奇异值变化为^[97]

$$\sigma_k^s = \sigma_k / \lambda^{3/2} \quad (3-72)$$

式中, σ_k 为原始图像 Radon 投影矩阵的第 k 个奇异值。由此可知, 当图像产生尺度变换时, Radon 投影矩阵的奇异值尺度变换因子为 $\lambda^{-3/2}$,



因此可以提取投影矩阵的奇异值对 Radon 矩特征进行规范化, 消除尺度变换引起中心距的影响。

令 $m = \max(\sigma_k^{2/3})$, $\eta_k(\theta) = \mu_k(\theta) / m^{2+k}$, 设 $\eta_k^s(\theta)$ 为图像产生尺度变换后的中心矩, 则

$$\begin{aligned}\eta_k^s(\theta) &= \mu_k^s(\theta) / (m^s)^{2+k} = \mu_k^s(\theta) / (\max((\sigma_k^s)^{2/3}))^{(2+k)} \\ &= \frac{\mu_k(\theta) / \lambda^{2+k}}{(\max((\sigma_k / \lambda^{3/2})^{2/3}))^{2+k}} = \frac{\mu_k(\theta) / \lambda^{2+k}}{(\max(\sigma_k^{2/3}) / \lambda)^{2+k}} = \frac{\mu_k(\theta) / \lambda^{2+k}}{(\max(\sigma_k^{2/3}))^{2+k} / \lambda^{2+k}} \\ &= \frac{\mu_k(\theta)}{(\max(\sigma_k^{2/3}))^{2+k}} = \mu_k(\theta) / m^{(2+k)} \\ &= \eta_k(\theta)\end{aligned}\tag{3-73}$$

在构造尺度不变性时, 也可以采用其他的参数进行尺度归一化 (如 \bar{t}_θ , $\mu_0(\theta)$ 等), 经多次重复实验表明采用 Radon 投影数据的奇异值作为比例因子消除图像尺度影响的效果较好, 这样 $\eta_k(\theta)$ 即为尺度和平移不变量。

3. 旋转不变量

由于矩阵的奇异值具有与其行列位置无关的特性, 根据该特性可获得旋转不变量特征。因此可以利用式 (3-73) 计算矩向量 $\eta = [\eta_2, \eta_3, \dots, \eta_k]$, 计算该不变矩的矩阵奇异值, 将得到的奇异值进行降序排列组成一向量, 该向量就是表示该图像具有平移、尺度和旋转不变性的特征向量。由于得到的奇异值范围差距较大, 在模式识别中往往采用前两项较大的特征作为特征向量, 而针对图像检索而言, 特征向量较少会导致分辨率较低, 因此对奇异值进行取对数, 然后取绝对值作为图像的特征向量, 即

$$\begin{cases} S = \text{SVD}(\eta) \\ S^* = \text{abs}(\log S) \end{cases}\tag{3-74}$$

得到的奇异值 S^* 则具有平移、尺度和旋转不变性。

为了评价算法的检索性能, 从 Corel 检索图像库选取了 1500 图像



作为检索测试库,包括恐龙、花卉、马、大象、公交车各类图像,相似图像均超过 30 幅。图 3-4 所示为 Radon 轮廓不变矩算法针对恐龙图像的检索结果。



图 3-4 Radon 轮廓不变矩算法针对恐龙图像的检索结果

3.3.5 基于显著封闭边界的形状特征

图像的边界可以表征形状特征,文献[98-99]等提出了基于显著边界的形状特征检索算法。首先,使用经典的 Canny 边缘检测算子对灰度图像进行边缘检测,获取边缘碎片图;其次,使用图论方法的最小边缘比算法提取图像中物体的显著封闭边界;然后,对显著封闭边界使用两个特征(边缘角度直方图和边缘幅度直方图)来描述,从而生成特征矢量。

1. 碎片的预处理

边缘检测算子检测到的边缘碎片并非全部都是互相断开的,往往会有碎片互相交叉、依附甚至封闭的情况出现,这种情况对图模型的构建是不利的。因此,为了解决这个特殊问题,需要断开这些连接以构建图



模型。首先,对于有交叉点的曲线(如图 3-5 (a) 所示),找出交叉点,并把它们断开以得到多个碎片(如图 3-5 (b) 所示),一个交叉点被断开为 3 个端点,在构建图模型时,它们(u_1, u_2, u_3)被具有零权重和零长度的点边缘连接。碎片依附是指两个碎片连接成一个碎片,由于这些依附的碎片可能会从图模型中排斥许多需要的点边缘,因此将大大损害显著边界检测的可靠性。对于碎片依附曲线(如图 3-5 (c) 所示),可以通过在高曲率处断开所有碎片,然后用具有零权重和零长度的点边缘连接 3 个端点来解决这个问题,如图 3-5 (d) 所示。对于封闭曲线(如图 3-5 (e) 所示),可以看做是依附的特例,简单地在最高曲率处断开封闭曲线,然后用具有零权重和零长度的点边缘连接两个端点即可,如图 3-5 (f) 所示。

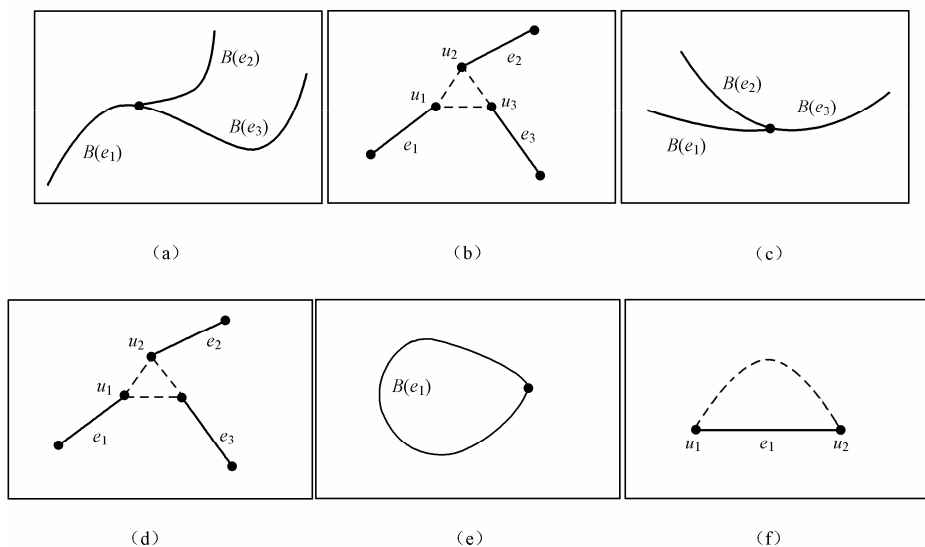


图 3-5 特殊情况下的碎片识别和图的构造

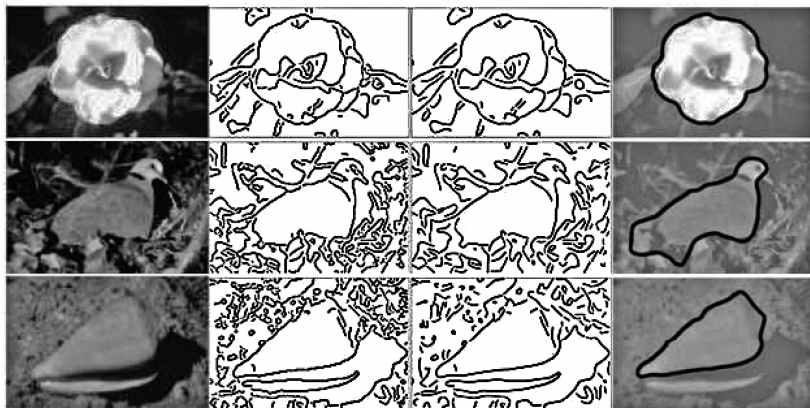
针对碎片进行预处理后,就可以检测碎片的平滑和缝隙的填充,从真实碎片中导出虚碎片,然后就可以进行显著封闭边界的提取。

2. 显著封闭边界的提取

检测出边缘碎片集后,对这些碎片集进行预处理,在这些预处理的



碎片集中选择一个子集，并把它们顺序地连接起来的过程称为边界抽取。为了反映图像的显著结构，抽取出来的边界是具有最大显著度的封闭边界，可以利用最小边缘比算法抽取的显著封闭边界，图 3-6 所示的是从自然图像中抽取的边界示意图。



(a) 原始图像 (b) Canny检测碎片 (c) 预处理后的碎片 (d) 算法抽取的显著封闭边界

图 3-6 从 3 幅自然图像抽取的边界

3. 显著封闭边界的特征描述

用两个特征来描述显著封闭边界。

1) 边缘幅度直方图 获得显著封闭边界后，用 Sobel 算子计算显著封闭边界的梯度幅度：

$$M(x, y) = \left[\left(\frac{\partial f(x, y)}{\partial x} \right)^2 + \left(\frac{\partial f(x, y)}{\partial y} \right)^2 \right]^{1/2} \quad (3-75)$$

将幅度图 M 分割为 6×6 的子块，统计每一个子块中显著封闭边界的个数，对统计结果生成一个直方图（每个子块中显著封闭边界特征点个数的取值范围是 $0 \sim 36$ ）。将归一化后的结果作为最终的幅度直方图。

2) 边缘角度直方图 获得显著封闭边界后，用 Sobel 算子计算显著封闭边界的梯度角度：

$$\theta(x, y) = \arg \tan \left[\left(\frac{\partial f(x, y)}{\partial x} \right) / \left(\frac{\partial f(x, y)}{\partial y} \right) \right] \quad (3-76)$$



由式 (3-76) 得到的角度取值范围为 $0^\circ \sim 360^\circ$ ，将它均匀量化为 36 个区间 (bins)，统计每个区间的像素数并进行归一化，生成归一化的角度直方图。对归一化后的直方图进行傅里叶变换使特征量具备旋转不变性，计算每一个变换分量的模值，将变换后的模值直方图作为图像显著封闭边界最终的梯度角度直方图特征矢量。

3.4 本章小结

本章首先介绍了形状特征概述，人类对于图像的理解在一定程度上依赖于图像中目标形状的感知，因此基于形状特征的图像检索是 CBIR 中重要的研究内容。目前，形状特征的描述主要分为基于区域的方法和基于边界的方法两大类，基于区域的方法是利用形状的全部区域研究，主要包括区域形状参数、HU 不变矩、Legendre 矩、Zernike 矩、Pseudo-Zernike 矩、Tchebichef 矩和 Krawtchouk 矩等方法；基于边界的方法主要有傅里叶形状描述符、基于内角的形状特征、小波 HU 不变矩、Radon 轮廓不变矩和基于显著封闭边界的形状特征。实验分析表明，基于区域分析的形状特征检索算法充分利用了图像的区域信息，具有较好的检索性能，但时间复杂度较高。

第四章



基于纹理特征的图像检索

4.1 纹理特征概述

纹理是重要的视觉特征之一，有关纹理的描述，目前尚无统一的定义。字典中，纹理被定义为由紧密地交织在一起的单元组成的某种结构。一般而言，纹理描述了图像所对应景物的表面特性，是对局部区域中像素之间关系的一种度量。从直观上看，纹理具有以下 3 个特征。

(1) 某种局部的有序性在该序列更大的区域内不断重复。

(2) 该序列由非随机排列的基本部分组成。

(3) 纹理区域内大都是均匀统一的统一体，图像特征显示出重复性结构。如木制家具上的木纹、花布上的花纹都属于纹理。

图像纹理在许多领域中得到了应用，如在场景分析、医学图像分析、遥感、天气预报和图像数据库检索等众多应用领域中都具有重要作用。目前，对纹理特征的研究方法是多种多样的，从分析技术角度归纳起来可以分为两大类，即基于空域的纹理分析和基于频域的纹理分析。基于空域的纹理分析通常利用图像的统计特征进行纹理分析，当纹理较细微时，统计方法较为适用，然而空域的纹理分析方法通常会受到噪声等影响，因此通常需要对纹理图像进行预处理。基于频域的纹理分析方法，借助傅里叶或小波频谱特性来描述周期性或近似周期性的纹理图像，该



方法将图像信号按照频率进行划分,得到不同频率的图像信号,有利于对图像信号进行去噪,可以得到稳定的纹理频域特征。下面将介绍空间域的纹理分析和频率域的纹理分析算法,同时分析了近年来的 3 种旋转不变性的纹理分析算法。

4.2 空间域的纹理分析

4.2.1 Tamura纹理特征

Tamura 等根据人类对纹理的视觉感知心理学的研究,提出了粗糙度 (Coarseness)、对比度 (Contrast)、方向度 (Directionality)、线像度 (Linelikeness)、规整度 (Regularity) 和粗略度 (Roughness) 等纹理特征的表达^[100]。Tamura 纹理特征的 6 个分量对应于心理学角度上纹理特征的 6 种属性。其中,粗糙度、对比度和方向度 3 个纹理特征在图像检索中具有较好的检索性能^[101],下面将重点讨论这 3 种视觉特征。

1. 粗糙度

粗糙度反映了图像纹理的粒度,其计算方法如下所述。

首先,计算图像中大小为 $2^k \times 2^k$ 个像素的活动窗口中像素的平均强度值,即

$$A_k(x, y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} g(i, j) / 2^{2k} \quad (4-1)$$

式中, $g(i, j)$ 是位于 (i, j) 的像素强度值, $k = 0, 1, \dots, 5$ 。然后,分别计算每个像素在水平和垂直方向上互不重叠的窗口之间的平均强度差,即

$$E_{k,h}(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)| \quad (4-2)$$

$$E_{k,v}(x, y) = |A_k(x, y + 2^{k-1}) - A_k(x, y - 2^{k-1})| \quad (4-3)$$

对于每个像素而言,能使 E 值达到最大 (无论方向) 的 k 值可以设



置最佳尺寸 $S_{\text{best}}(x, y) = 2^k$ 。最后, 粗糙度可以通过计算整幅图像中 S_{best} 的平均值来得到, 即

$$F_{\text{crs}} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n S_{\text{best}}(i, j) \quad (4-4)$$

式中, m 、 n 分别为图像的宽度和高度。

2. 对比度

对比度是反映视觉像素亮度的一个统计值, 它是通过 $\alpha_4 = \mu_4 / \sigma^4$ 来定义的, 其中 μ_4 是 4 次矩而 σ^2 是方差, 对比度计算公式为

$$F_{\text{con}} = \frac{\sigma}{\alpha_4^{1/4}} \quad (4-5)$$

该值给出了整个图像或区域中对比度的全局度量。

3. 方向度

方向度分析了图像纹理沿着某个方向分布是否集中或分散的统计值。首先需要计算每个像素处的梯度向量, 向量的模和方向分别定义为

$$|\Delta G| = (|\Delta_H| + |\Delta_V|) / 2 \quad (4-6)$$

$$\theta = \arg \tan(\Delta_V / \Delta_H) + \pi/2 \quad (4-7)$$

式中, Δ_H 是图像卷积式 (4-8) 下列两个 3×3 操作符所得的水平变化量, 而 Δ_V 是图像卷积式 (4-8) 下列两个 3×3 操作符所得的垂直方向上的变化量。

$$\begin{array}{ccc} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{array} \quad \begin{array}{ccc} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{array} \quad (4-8)$$

当所有像素的梯度向量都被计算出来后, 直方图 H_θ 被构造出来用于表达 θ 值。直方图对于具有明显方向性的图像会表现出峰值, 对于无明显方向的图像则表现得比较平坦。图像纹理的全局方向性可以通过直方图峰值的尖锐程度进行计算, 即



$$F_{\text{dir}} = \sum_p \sum_{\phi \in W_p}^{n_p} (\phi - \phi_p)^2 H_D(\phi) \quad (4-9)$$

式中, p 代表直方图中的峰值, n_p 为直方图中所有的峰值, W_p 代表某个峰值 p 所包含的所有的 bin, 而 ϕ_p 是具有最高值的 bin。

4.2.2 灰度共生矩阵法

灰度共生矩阵^[102]是图像灰度变化的二阶统计度量, 根据两个像素点位置的联合概率密度函数建立纹理图像的共生矩阵, 表述了图像纹理结构性统计特征。由于纹理是由灰度分布在空间位置上反复出现而形成的, 因而在图像空间中相隔某距离的两像素之间会存在一定的灰度关系, 即图像中灰度的空间相关特性。

灰度共生矩阵可表示为 $H = [p_\delta(i, j)]$, 其中 $\delta = (\Delta x, \Delta y)$ 表示纹理图像两像素点间的距离, 两个像素在 x 方向和 y 方向上距离分别为 Δx 和 Δy 。不同的 $\delta = (\Delta x, \Delta y)$ 决定了两像素间的距离和方向, 0° 方向为 $\delta = (0, \pm|\Delta y|)$; 90° 方向为 $\delta = (\pm|\Delta x|, 0)$; 45° 方向为 $\delta = (|\Delta x|, -|\Delta y|)$ 或 $\delta = (-|\Delta x|, |\Delta y|)$ 且 $|\Delta x| = |\Delta y|$; 135° 方向为 $\delta = (|\Delta x|, |\Delta y|)$ 或 $\delta = (-|\Delta x|, -|\Delta y|)$ 且 $|\Delta x| = |\Delta y|$ 。常用以上的 4 个方向的位置关系计算各自的灰度共生矩阵。

共生矩阵反映了图像灰度分布关于方向、局部邻域和变化幅度的综合信息, 但它不能直接提供区别纹理的特性。为了能描述纹理的性质, 可以提取共生矩阵的能量、对比度、相关性、信息熵和局部均匀性等, 下面是它们的定义。

➤ 能量(角二阶矩):

$$Q_1 = \sum_i \sum_j [p_\delta(i, j)]^2 \quad (4-10)$$

➤ 对比度

$$Q_2 = \sum_k k^2 [\sum_i \sum_j p_\delta(i, j)], (k = i - j) \quad (4-11)$$

➤ 相关性



$$Q_3 = \frac{\sum_i \sum_j ij p_\delta(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (4-12)$$

式中,

$$\mu_x = \sum_i i \sum_j p_\delta(i, j) \quad (4-13)$$

$$\mu_y = \sum_j j \sum_i p_\delta(i, j) \quad (4-14)$$

$$\sigma_x^2 = \sum_i (i - \mu_x)^2 \sum_j p_\delta(i, j) \quad (4-15)$$

$$\sigma_y^2 = \sum_j (j - \mu_y)^2 \sum_i p_\delta(i, j) \quad (4-16)$$

➤ 信息熵

$$Q_4 = - \sum_i \sum_j p_\delta(i, j) \log p_\delta(i, j) \quad (4-17)$$

➤ 局部均匀性

$$Q_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p_\delta(i, j) \quad (4-18)$$

4.2.3 LBP算法

为了获取局部纹理分布, Ojala^[103]提出了局部二值模式 (Local Binary Pattern, LBP) 表征图像的纹理基元, 该算法具有较好的检索效果。局部二值模式 LBP 是在二值的情况下, 将原始的 3×3 邻域通过中心像素取阈值来实现二值模式, 然后在阈值区域将像素值乘以相应的像素的权值获得最后的 LBP 模式, 如图 4-1 所示, 将邻域像素的灰度值表示为 g_i ($1 \leq i \leq 8$), 对于图像的任意像素, 根据八邻域的像素灰度值来计算其 LBP 值^[103-105]。

g_1	g_2	g_3
g_4	g_0	g_5
g_6	g_7	g_8

图 4-1 LBP 算法中的 3×3 邻域



邻域像素根据中心像素灰度值进行比较, 获得一个二值数值, 计算公式为

$$s(g_0, g_i) = \begin{cases} 1 & g_i \geq g_0 \\ 0 & g_i < g_0 \end{cases} (1 \leq i \leq 8) \quad (4-19)$$

将 8 个邻域像素的二值数值, 依照顺序编码求出中心像素的 LBP 值, 其计算方法为

$$LBP = \sum_{i=1}^8 s(g_0, g_i) 2^{i-1} \quad (4-20)$$

LBP 直方图计算方法如图 4-2 所示, 纹理直方图特征反映了物体的纹理分布。LBP 纹理直方图的计算方法如下所述。

- (1) 采用方差归一化方法处理灰度图像以消除光照影响。
- (2) 使用基本类型 LBP 算子把图像转换成 LBP 图像。
- (3) 求出 LBP 图像的直方图。

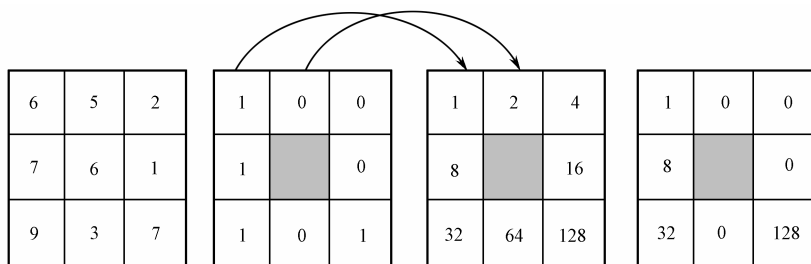


图 4-2 LBP 直方图计算方法

由图 4-2 可知, $LBP=1+8+32+128=169$ 。2002 年, Ojala^[105]提出了具有旋转和尺度不变的 LBP 算法, 设单色调图像局部邻域的纹理特征为 T , P ($P>1$) 个图像像素的灰度联合分布为

$$T = t(g_c, g_0, \dots, g_{p-1}) \quad (4-21)$$

其中 g_c 灰度值是局部邻域的中心值, g_p 是在半径为 R ($R>0$) 的圆形对称区域内 p 个相等空间像素灰度值。如果 g_c 的坐标是 $(0,0)$, 则 g_p 的坐标是 $(-R \sin(2\pi p / p), R \cos(2\pi p / p))$ 。为了消除尺度变化的影响, 可以通过只考虑差值的符号得到一个尺度不变性的算子, 式 (4-21) 可变为



$$T \approx t(s(g_0 - g_c), s(g_1 - g_c), \dots, s(g_{p-1} - g_c)) \quad (4-22)$$

式中, $s(x)$ 的定义为

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (4-23)$$

此时得到的分布中所有值都为 0 或 1, 这个局部结构可以看做是纹理单元。模板可以进一步扩大到邻域不同的尺寸, 应用循环邻域关系和像素的双线性内插可以使得邻域有不同的半径和不同的像素数。对于邻域关系这里应用符号 (P, R) , 其中 R 是采样半径, P 是采样点数, 统一模式的 LBP 模式具体的定义为^[104-105]

$$\text{LBP}_{P,R}^{\text{riu2}} = \begin{cases} \sum_{p=0}^{p-1} s(g_p - g_c) & \text{若 } U(\text{LBP}_{P,R}) \leq 2 \\ p+1 & \text{否则} \end{cases} \quad (4-24)$$

$$U(\text{LBP}_{P,R}) = |s(g_{p+1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (4-25)$$

其中上角的 riu2 反映了旋转不变量统一模式, U 的最大值为 2。图 4-3 所示为不同的 (P, R) 的圆形对称邻域集示意图。

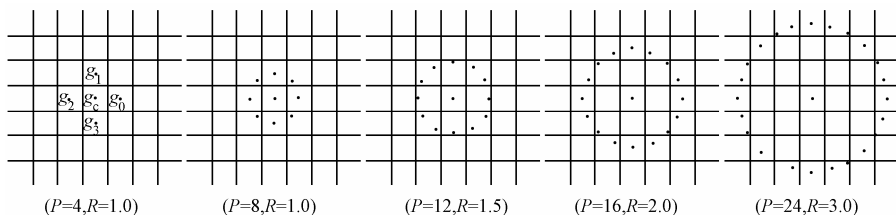


图 4-3 不同的 (P, R) 的圆形对称邻域集

Ojala 的实验表明, 使用尺度为 $(8, 1)$ 的 LBP 算子时, 统一模式的数量为 59, 占 LBP 算子总数的 23%, 但可以表达出其中 87.2% 的纹理, 因此 LBP 算子可以有效地描述图像的大部分纹理特征, 具有较好的检索性能^[2]。

4.2.4 Laws 纹理能量

纹理能量法是 1980 年由 Laws^[106]提出的一种纹理滤波分析法, 其



主要思想是针对图像进行直方图均衡化处理,在 3×3 或 5×5 窗口内用模板进行微观滤波计算,取卷积之和作为新图像输出,然后将卷积后的图像中每个像素值用以该像素为中心的 15×15 局部窗口内的均方差或绝对值平均代替,Laws 称窗口内的均方值度量为 Laws 纹理能量法。设 L 为代表层, E 表示边缘, S 表示形状, W 表示波, R 表示纹路,Laws 所有滤波模板出自以下 8 个最基本的一维算子,

$$L_3 = [1 \quad 2 \quad 1]$$

$$E_3 = [-1 \quad 0 \quad 1]$$

$$S_3 = [-1 \quad 2 \quad -1]$$

$$L_5 = [1 \quad 4 \quad 6 \quad 4 \quad 1]$$

$$E_5 = [-1 \quad -2 \quad 0 \quad 2 \quad 1]$$

$$S_5 = [-1 \quad 0 \quad 2 \quad 0 \quad -1]$$

$$W_5 = [-1 \quad 2 \quad 0 \quad -2 \quad 1]$$

$$R_5 = [1 \quad -4 \quad 6 \quad -4 \quad 1]$$

通过 L_3 、 E_3 和 S_3 三个基本算子相互卷积,以及通过 5 个 L_5 、 E_5 、 S_5 、 W_5 和 R_5 基本算子相互卷积可获得各种模板滤波器^[106-107]。2D 的模板可以由两个 1-D 模板进行外积获得,Laws 将 Brodatz 的 8 种纹理图像拼在一起,对图像做纹理能量测量,将每个像元指定为 8 个可能类中的一个,正确率达到 87%。实验分析表明,图 4-4 所示的 4 个 5×5 模板是最有效的算子。

-1	-4	-6	-4	-1
-2	-8	-12	-8	-2
0	0	0	0	0
2	8	12	8	2
1	4	6	4	1

(a) $L_5 E_5$ 水平边缘

-1	-4	-6	-4	-1
-4	16	-24	16	-4
-6	-24	36	-24	-6
-4	16	-24	16	-4
-1	-4	-6	-4	-1

(b) $R_5 R_5$ 高频斑点

图 4-4 四个有效的 Laws 模板滤波器



-1	0	2	0	-1
-2	0	4	0	-2
0	0	0	0	0
2	0	-4	0	2
1	0	-2	0	1

(c) $S_5 E_5$

-1	0	2	0	-1
-4	0	8	0	-4
-6	0	12	0	-6
-4	0	8	0	-4
-1	0	2	0	-1

(d) $L_5 E_5$ 垂直线

图 4-4 四个有效的 Laws 模板滤波器 (续)

Laws 纹理能量法简单有效,但是所提供的模板较少,没有较多的变化性质,因此该方法的应用受到一定局限。

4.3 频率域纹理分析

4.3.1 傅里叶功率谱法

傅里叶变换原理表明,任何连续测量的时序或信号,都可以表示为不同频率的正弦波信号的无限叠加,以累加方式来计算该信号中不同正弦波信号的频率、振幅和相位。傅里叶变换将原来难以处理的时域信号转换成了易于分析的频域信号(信号的频谱),可以利用一些工具对这些频域信号进行处理、加工。

傅里叶变换具有如下性质:①傅里叶变换是线性算子,若赋予适当的范数,它还是酉算子;②傅里叶变换的逆变换容易求出;③正弦基函数是微分运算的本征函数,从而使得线性微分方程的求解可以转化为常系数的代数方程的求解;④卷积定理指出,傅里叶变换可以化复杂的卷积运算为简单的乘积运算,从而提供了计算卷积的一种简单手段;⑤离散形式的傅里叶变换可以利用快速傅里叶变换算法 FFT 算出。

傅里叶功率谱法是描述图像纹理特征的一种频率域方法,设图像为 $f(x, y)$, 其定义域为 Φ , $f(x, y)$ 的傅里叶变换可由下式来表示:



$$F(u, v) = \iint_{\phi} f(x, y) \exp\{-j2\pi(ux + vy)\} dx dy \quad (4-26)$$

一般情况, $F(u, v)$ 是一个复数, 因此二维傅里叶变换的功率谱可写成:

$$|F(u, v)|^2 = F(u, v)F^*(u, v) \quad (4-27)$$

式中, F^* 是 F 的共轭, $|F(u, v)|$ 是一个实数。为了讨论简单, 将式 (4-27) 用极坐标的形式来表示, 记为 $F(r, \theta)$, $|F(r, \theta)|^2$ 沿径向的分布可以作为子图像区域 ϕ 的局部纹理结构的一种度量。如果不考虑纹理取向, 则纹理的粗糙性的度量为

$$R(r) = \int_0^{2\pi} [F(r, \theta)]^2 d\theta \quad (4-28)$$

对式 (4-28) 取不同的 r 值, 可以得到区域 ϕ 的一组纹理结构特性。实验证明, 仅利用功率谱作为图像纹理特征, 其检索效果较差^[108-109]。

4.3.2 小波变换统计矩

小波分析是在短时傅里叶变换的基础上发展起来的一种新的时频分析方法, 和傅里叶分析相比有着许多本质上的进步。小波变换通过伸缩和平移等运算功能对函数或信号进行多尺度细化分析, 其主要特点是具有对时频域双重局部定域能力和多分辨分析能力, 被誉为“数学显微镜”, 能够有效地从信号中提取信息, 已成为瞬变信号分析的有力工具^[110]。

1910 年, A.Haar 提出第一个小波规范正交基, 即 Haar 小波, 不过当时还没有“小波”概念。1981 年, 法国地质物理学家 J.Moriet^[111]在分析地质数据时, 基于群论首先提出了小波分析的概念。1988 年, Daubechies 构造了紧支撑标准正交小波基^[112], 即一系列具有任意选定的正则性、有限支撑、正交的尺度函数和小波, 并将当时所有正交小波的构造统一起来。1989 年, S.Mallat^[113]将计算机视觉领域内多尺度分析的思想引入到小波分析中, 提出多分辨分析 (Multiresolution Analysis) 的概念, 用多分辨分析来定义小波变换, 提出快速小波变换算法 (即



Mallat 算法)，并将它用于图像分析和完全重构。

1-D 小波变换是将信号分解为一系列的基本函数 $\psi_{mn}(x)$ ，这些基本函数都是通过对母函数 $\psi(x)$ 的变形得到，即

$$\psi_{mn}(x) = 2^{-m/2} \psi(2^{-m}x - n) \quad (4-29)$$

式中， m 和 n 是整数。这样，信号 $g(\beta)$ 可以被表达为

$$f(x) = \sum_{mn} c_{mn} \psi_{mn}(x) \quad (4-30)$$

二维小波变换的计算需要进行递归地过滤和采样。在每个层次上，二维的信号被分解为 4 个子波段，根据频率特征分别称为 LL、LH、HL 和 HH。设 $\phi(x, y)$ 是一个可分离的二维尺度函数，即

$$\phi(x, y) = \phi(x)\phi(y) \quad (4-31)$$

式中， $\phi(x), \phi(y)$ 是一维尺度函数，若 $\psi(x), \psi(y)$ 是相应的小波，则可确定下列 3 个二维基本小波：

$$\psi^1(x, y) = \phi(x)\psi(y) \quad (4-32)$$

$$\psi^2(x, y) = \psi(x)\phi(y) \quad (4-33)$$

$$\psi^3(x, y) = \psi(x)\psi(y) \quad (4-34)$$

根据二维小波，图像的二维离散小波变换为

$$\left. \begin{aligned} f_{2^{j+1}}^0(m, n) &= \{[f_{2^j}^0(x, y) * \phi(-x, -y)]\}(2m, 2n) \\ f_{2^{j+1}}^1(m, n) &= \{[f_{2^j}^0(x, y) * \psi^1(-x, -y)]\}(2m, 2n) \\ f_{2^{j+1}}^2(m, n) &= \{[f_{2^j}^0(x, y) * \psi^2(-x, -y)]\}(2m, 2n) \\ f_{2^{j+1}}^3(m, n) &= \{[f_{2^j}^0(x, y) * \psi^3(-x, -y)]\}(2m, 2n) \end{aligned} \right\} \quad (4-35)$$

式中， $f_{2^0}^0(x, y)$ 是原图像，经过离散小波变换，图像分解为 $1/4$ 大小的子图：水平方向、垂直方向、对角线方向的高频部分细节图像和低频部分近似图像，每个图像都按“二抽取”滤波得到。近似图像 $f_{2^{j+1}}^0(x, y)$ 以相同的方式再分解成 2^{j+1} 尺度上的更小的图像，原始图像就被多分辨率分解成不同空间和不同频率的多个子图^[99]。图像的离散小波分解的二层分解如图 4-5 所示。

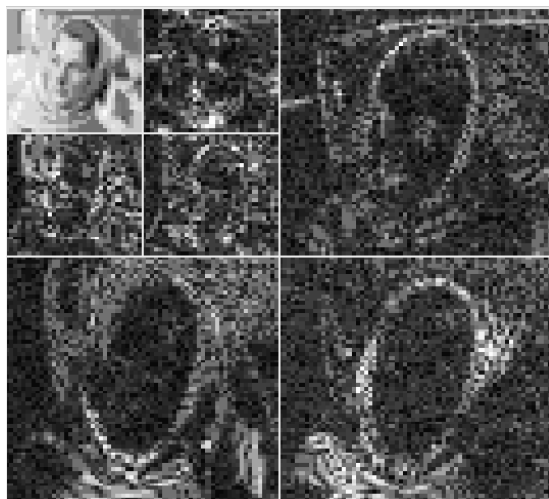


图 4-5 图像的离散小波分解的二层分解

本节采用三层小波分解，共有 10 个子带，利用各子带系数能量的均值和方差作为图像纹理特征，子带 j ($j=1,2,\dots,10$) 的均值和方差可以定义为

$$m_1 = \frac{1}{N^2} \sum_{x=1}^N \sum_{y=1}^N I(x, y) \quad (4-36)$$

$$\text{Std} = \left(\frac{1}{N^2 - 1} \sum_{x=1}^N \sum_{y=1}^N (|I(x, y)| - m_1)^2 \right)^{1/2} \quad (4-37)$$

经过小波变换的分解，图像的能量分布在不同的子带内，低频能量主要集中在 LL 子带中，而高频能量分布于其他子带中。下面计算小波变换域常见的 4 种能量统计特征，则第 j 个子带 $I(x, y)$ 的能量可以表示为

$$E_1 = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N |I(x, y)| \quad (4-38)$$

$$E_2 = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N I^2(x, y) \quad (4-39)$$

$$E_3 = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N \sqrt{I(x, y)} \quad (4-40)$$



$$E_4 = \frac{-1}{\log(M \times N)} \sum_{x=1}^M \sum_{y=1}^N \frac{\sqrt{I(x,y)}}{\text{snorm}} \quad (4-41)$$

式中, $\text{snorm} = \sum_x \sum_y \sqrt{I(x,y)}$ 。实验分析表明, 同时采用 E_1 和 E_2 能量特征进行检索具有较好的检索性能。

4.3.3 小波直方图

小波变换子带系数的均值、方差和能量统计特征表征了每个子带的统计信息, 而没有表征各子带系数之间的系数联系, 因此 Smith^[18]提出了一种小波直方图 (Wavelet Histogram Technique, WHT) 技术, WHT 考虑了各子带系数之间的联系, 可以用做图像数据库的索引。其实现过程为, 首先将图像进行 3 层离散小波变换 (DWT) 产生 10 个小波子带; 其次, 取 9 个高通子带小波系数的绝对值, 并对 9 个子带进行“过采样”, 将子带扩充为原图像大小, 根据 9 个纹理通道的值, 一个纹理点可表示为一个 9 维向量, 9 维向量的每个元素都是二进制值 0 或 1; 最后, 用直方图表示各个纹理向量, 这样就得到了小波直方图。小波直方图的产生框架如图 4-6 所示。

小波变换直方图在一定程度上表征了图像的纹理特征, 然而其检索性能并不理想, 同时高频子带系数的统计信息没有充分利用。从小波直方图技术的产生过程可以看出, 它的计算复杂性是很高的, 这主要是由于将 9 个高通子带“过采样”成原图像大小和阈值化造成的。同时, 由于只使用 9 个高通子带, 没有使用低通子带 A_0 , 故 WHT 用于检索纹理图像的效果较好, 对检索自然图像效果较差。Mandal^[114]在 WHT 的基础上提出了快速小波直方图技术 (Fast Wavelet Histogram Techniques, FWHT), FWHT 算法在一定程度上减少了计算复杂度, 同时 Mandal 提出了 3 种 FWHT 的特征提取方案, 其中图 4-6 可以作为一种 FWHT 方案, 但使用了第 10 个子带, 图 4-7 和图 4-8 所示的是其中的两种方案。

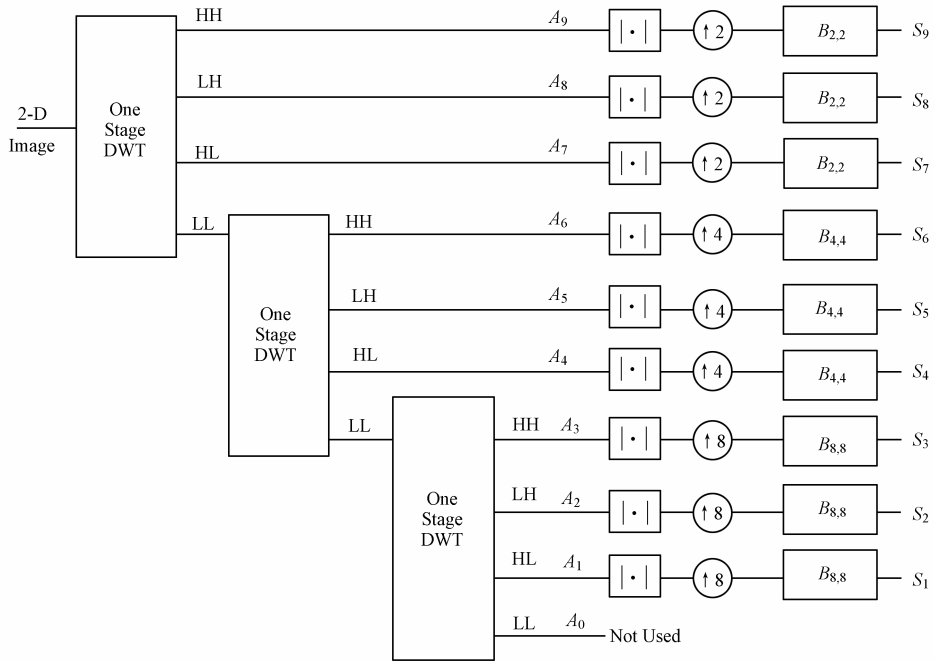


图 4-6 小波直方图的产生框架

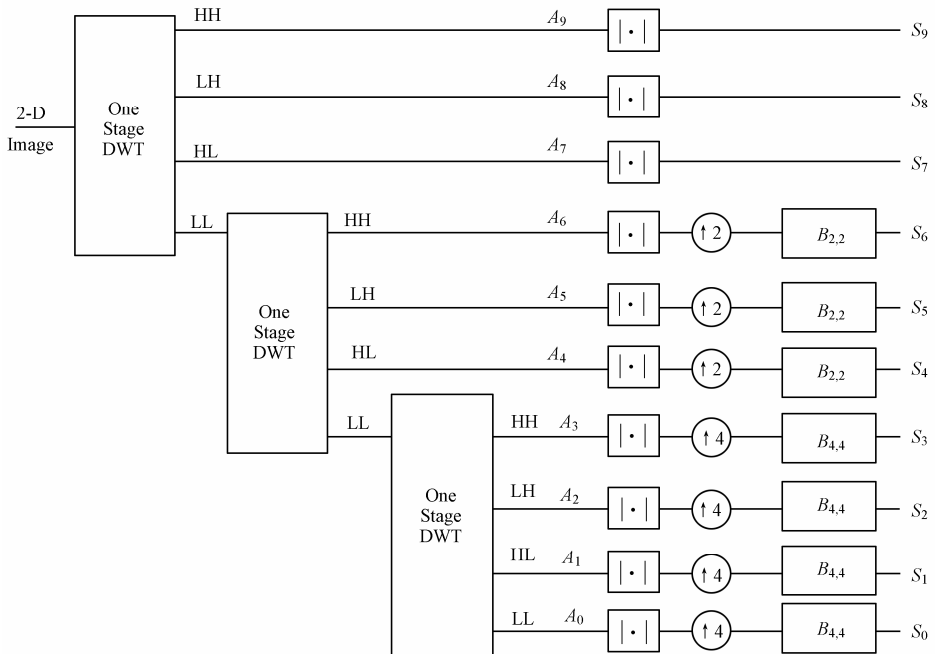


图 4-7 快速小波直方图方案 1 框架

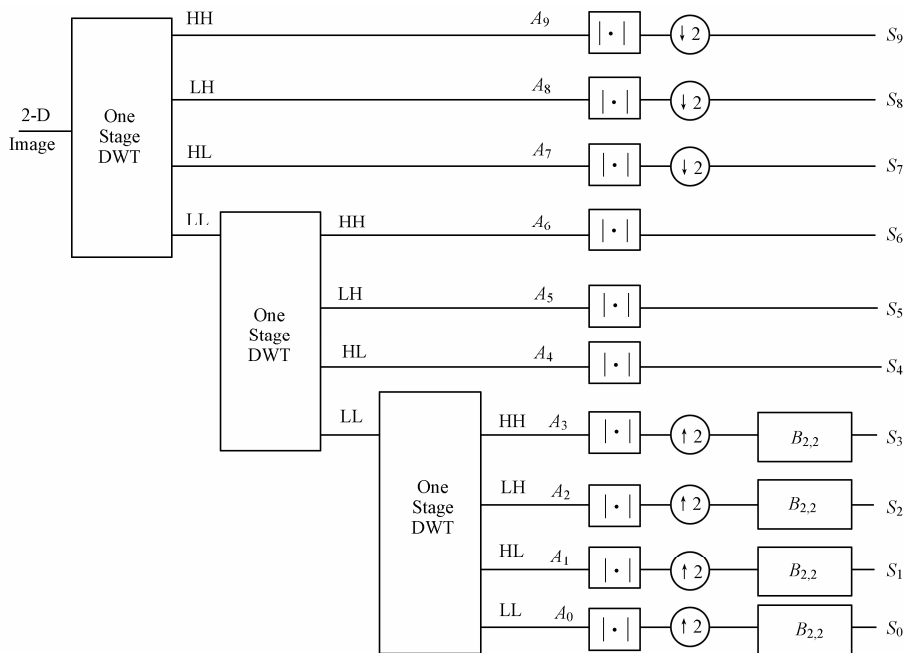


图 4-8 快速小波直方图方案 2 框架

基于快速小波直方图的 3 种方案检索试验表明, 方案 1 的检索算法检索性能最高, 但计算复杂度非常高。为了进一步降低快速小波直方图算法, Smith 等^[18]提出可以采用低频 7 个子带进行直方图提取, 从而大大降低了检索算法的时间复杂度, 实验证明其检索性能在一定程度上有一些降低, 但差距不大。

4.3.4 小波GGD&KLD

Minh^[28]分析了小波变换高频子带系数的直方图, 分析表明小波系数直方图具有非高斯性、高峰度和长拖尾等特点, 可以采用广义高斯模型 (Generalized Gaussian Density, GGD) 进行描述。小波系数直方图本质上是小波子带系数的边缘分布概率密度图, 广义高斯密度模型表征了小波系数的直方图分布。实验表明, 广义高斯密度函数可通过改变两个参数 α , β 来很好地逼近各种小波变换所得的各小波子带系数的边缘分布概率密度函数。广义高斯密度函数定义为^[28,115]



$$p(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta} \quad (4-42)$$

其中 $\Gamma(\cdot)$ 为 Gamma 函数, 即 $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt \quad t > 0$, α 表征了概率密度函数的尖峰宽度 (标准差), β 反比于尖峰的下降速度, 因此 α , β 分别被称为尺度参数和形状参数, 同时 Minh 提出采用极大似然法估计以上两个参数。设小波高频子带系数的直方图为 $x = (x_1, x_2, \dots, x_L)$, 则似然函数为

$$L(x; \alpha, \beta) = \log \prod_{i=1}^L p(x_i; \alpha, \beta) \quad (4-43)$$

针对 α , β 参数进行极大似然估计可得,

$$\frac{\partial L(x; \alpha, \beta)}{\partial \alpha} = -\frac{L}{\alpha} + \sum_{i=1}^L \frac{\beta |x_i|^\beta \alpha^{-\beta}}{\alpha} = 0 \quad (4-44)$$

$$\frac{\partial L(x; \alpha, \beta)}{\alpha \beta} = \frac{L}{\beta} + \frac{L\psi(1/\beta)}{\beta^2} - \sum_{i=1}^L \left(\frac{|x_i|}{\alpha} \right)^\beta \log \left(\frac{|x_i|}{\alpha} \right) = 0 \quad (4-45)$$

当固定 $\beta > 0$ 后, 可得,

$$\hat{\alpha} = \left(\frac{\beta}{L} \sum_{i=1}^L |x_i|^\beta \right)^{1/\beta} \quad (4-46)$$

将式 (4-46) 代入式 (4-45) 中得,

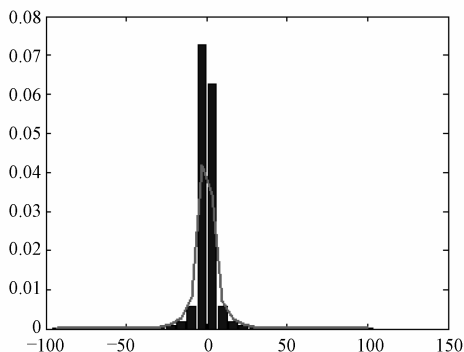
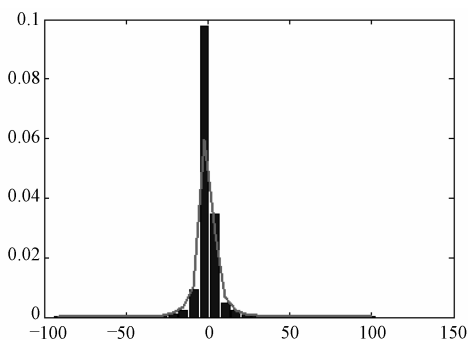
$$1 + \frac{\psi(1/\hat{\beta})}{\hat{\beta}} - \frac{\sum_{i=1}^L |x_i|^{\hat{\beta}} \log |x_i|}{\sum_{i=1}^L |x_i|^{\hat{\beta}}} + \frac{\log \left(\frac{\hat{\beta}}{L} \sum_{i=1}^L |x_i|^{\hat{\beta}} \right)}{\hat{\beta}} = 0 \quad (4-47)$$

利用牛顿—辛普森方法迭代求解出 $\hat{\beta}$ 。Minh 提出采用 KLD 距离测度表征两广义高斯密度函数间相似度。实验结果表明, 采用 GGD&KLD 距离的检索算法要优于小波变换的统计矩方法。

图 4-9 所示为辣椒图像, 图 4-10 所示的是图 4-9 的单层小波变换水平细节 CH1 的 GGD 拟合示意图, 其中红色是采用极大似然法的 GGD 拟合曲线; 图 4-11 所示的是垂直细节 CV1 的 GGD 拟合示意图, 图 4-12 所示的是对角细节 CD1 的 GGD 拟合示意图。



图 4-9 辣椒图像

图 4-10 水平细节 CH1 的 GGD 拟合 ($\alpha=0.3063$, $\beta=0.4256$)图 4-11 垂直细节 CV1 的 GGD 拟合 ($\alpha=0.3803$, $\beta=0.4410$)

由于小波变换的 GGD&KLD 算法具有较好的检索性能, 尚赵伟等^[116]利用拟合度实验分析了双密度小波变换的子带分解, 分析表明双密度小波分解高频子带符合广义高斯分布, 同时将双密度小波变换进行组合, 计算组合后的子带信息的 GGD。实验分析表明, 组合双密度小



波变换的 GGD&KLD 算法检索性能优于传统的小波变换。

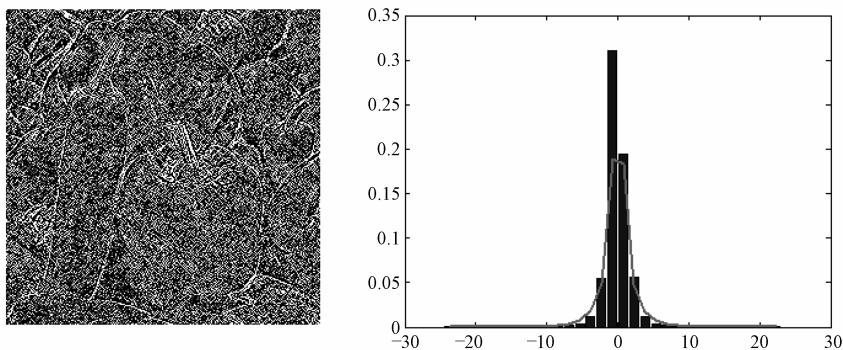


图 4-12 对角细节 CD1 的 GGD 拟合 ($\alpha=0.3652$, $\beta=0.5750$)

4.3.5 Gabor小波纹理分析

Gabor 变换具有多通道和多分辨率的特性,可同时在空间域和频率域达到最优联合分辨率,因此已被广泛应用于纹理分析和图像识别中。在纹理分析中,常用的 Gabor 函数变换对为^[117]

$$\begin{cases} g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi j W_x\right] \\ G(u, v) = \exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right]\right\} \end{cases} \quad (4-48)$$

式中,

$$\sigma_u = \frac{1}{2\pi\sigma_x}, \sigma_v = \frac{1}{2\pi\sigma_y} \quad (4-49)$$

Gabor 函数集构成了一组完备非正交基,使用这组基来描述一个信号可以得到一个局部化的频率描述。假设 $g(x,y)$ 是 Gabor 小波的母函数,则 Gabor 滤波器组就可以通过适当的旋转和缩放 $g(x,y)$ 得到:



$$\begin{cases} g_{mn}(x, y) = a^{-m} g(x', y') & a > 1, m, n = 0, 1, 2, \dots \\ x' = a^{-m} (x \cos \theta + y \sin \theta) \\ y' = a^{-m} (-x \sin \theta + y \cos \theta) \\ \theta = n\pi / K \end{cases} \quad (4-50)$$

Gabor 滤波器组的设计就依靠 a, σ_x, σ_y 这 3 个参数的设定。Gabor 小波的非正交性说明有冗余，因此需要尽量减少这种冗余。设 m 是 Gabor 滤波器的方向数， n 代表伸缩尺度数，文献[109]中的策略就是使得滤波器在频率响应时，在半峰幅值处（Half-peak magnitude）彼此接触，如图 4-13 所示。

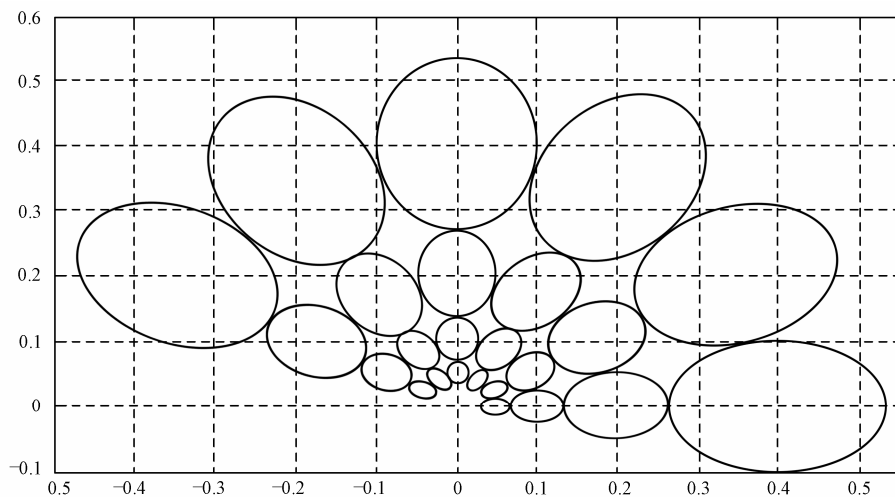


图 4-13 六个方向、四个尺度组成的 Gabor 滤波器组在频率域中响应的半峰幅值

具体的参数为

$$\begin{cases} a = (U_h / U_l)^{\frac{1}{s-1}} \\ \sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}} \\ \sigma_v = \tan\left(\frac{\pi}{2k}\right) \left[U_h - 2 \ln\left(\frac{\sigma_u^2}{U_h}\right) \right] \left[2 \ln 2 - \frac{(2 \ln 2)^2 \sigma_u^2}{U_h^2} \right]^{-\frac{1}{2}} \\ W = U_h \quad m = 0, 1, \dots, S-1 \end{cases} \quad (4-51)$$



在应用 Gabor 小波进行纹理分析和图像检索时,可以利用子带的均值和方差表征图像的纹理特征,即

$$\begin{cases} \mu_{mn} = \iint |W_{mn}(x, y)| dx dy \\ \sigma_{mn} = \sqrt{\iint (|W_{mn}(x, y)| - \mu_{mn})^2 dx dy} \end{cases} \quad (4-52)$$

实验分析表明,利用 Gabor 小波进行图像检索符合人类的视觉感知系统,具有较高的检索性能。近年来, Moghaddam^[118]采用 Gabor 小波变换的自相关系数作为图像纹理描述算子,实验表明优于传统小波变换的自相关系数算法。

4.3.6 多小波变换纹理分析

目前,基于频率域的小波检索算法基本上采用了单小波提取图像的特征,然而单小波不能够同时具备正交、对称、紧支撑和高阶消失矩等特性,在一定程度上限制了这些算法的检索效果。针对已有工作的一些不足之处,尚赵伟^[120]提出了基于多小波变换的纹理图像检索。

1. 多小波变换

多小波^[119]是指由两个或两个以上的函数作为尺度函数生成的小波,由于多小波构造具有较大的自由度,同时具备正交、对称、短紧支撑和高消失矩等特性,因此理论上多小波比单小波具有更多的优势。多小波变换采用了多滤波器,与传统的单滤波器不同,它是一个多输入、多输出系统,信号在多滤波前必须解决好向量输入流的生成和多尺度函数。多小波分量导致谱行为带来的不平衡通道问题,所以在多滤波前需进行预滤波处理,保证后续处理的正确性。数据经过预滤波处理后,产生的向量数据流通过低通、高通滤波器生成两个向量输入流,通过对生成的向量输入流进行相应处理,可以实现指定的功能。在逆多小波变换后,还需将向量数据流还原成与之对应的标量数据流,即需要进行后置滤波。多小波的定义如下所述。



【定义】设 $\phi(x) = [\phi_1, \phi_2, \dots, \phi_r]^T$ 是一个向量值函数, 其中 $\phi_i \in L^2(R)$, $i=1, 2, \dots, r$ 。对于 $j \in Z$, $V_j = \overline{\text{span}\{2^{j/2} \phi_i(2^j x - k) : 1 \leq i \leq r, k \in Z\}}$, 闭空间序列 $\{V_j\}_{j \in Z}$ 称为形成一个多分辨分析, 如果它满足如下条件。

- (1) 单调性: $\subset V_0 \subset V_1 \subset V_2 \subset \dots$
- (2) 逼近性: $\overline{\bigcup V_j} = L^2(R)$; $\bigcap V_j = \{0\}$;
- (3) 伸缩性: $f \in V_j \Leftrightarrow f(2x) \in V_{j+1}$;
- (4) 平移不变性: $f \in V_j \Leftrightarrow f(x-k) \in V_j, \forall k \in Z$;
- (5) Riesz 基存在性: $\{\phi_i(2^j x - k) : 1 \leq i \leq r, k \in Z\}$ 构成 V_j 子空间的 Riesz 基。

在上述情况下, $\phi(x)$ 称为 r 重尺度函数。定义 W_j 是 V_j 在 V_{j-1} 的正交补空间, 向量函数 $\psi(x) = [\psi_1, \psi_2, \dots, \psi_r]^T$, $\psi_i \in L^2, i=1, 2, \dots, r$, 其伸缩平移构成 W_j 子空间的一个 Riesz 基, 即

$$W_j = \overline{\text{span}\{2^{j/2} \psi_i(2^j x - k) : 1 \leq i \leq r, k \in Z\}} \quad (4-53)$$

此时, $\psi(x)$ 称为多小波函数。图像数据是二维数据, 可以利用分离法, 通过两次一维多小波变换实现二维多小波变换。

设用 $N \times N$ 矩阵 I_0 表示图像数据, 则分解过程如下所述: 对 I_0 中的每一行做预处理的 I_1 , I_1 中的每一行的前一半数据为与第一个尺度函数对应的系数, 后一半数据为与第二个尺度函数对应的系数; 对 I_1 中的每一列进行预处理的 I_2 , I_2 每一列的前一半数据为与第一个尺度函数对应的系数, 后一半数据为与第二个尺度函数对应的系数; 对 I_2 中的每一行进行一维的多小波变换得 I_3 ; 对 I_3 中的每一列进行一维多小波变换得到 I_4 。以上 4 步完成了一级二维多小波分解。若进行多级多小波分解, 只需将低频部分继续进行分解即可。

单小波只有一个尺度函数, 变换后的每个子带只有一个子图, 而多小波有 r 个尺度函数, 因此变换后每个子带有 $r \times r$ 个子图。容易证明, L 级多小波变换将图像分解为 $r^2 \times (3L+1)$ 个子图^[120]。图 4-14 所示为三层多小波分解示意图, 其中 $r=2$, 三层多小波分解共有 10 个子带,



而每个子带有 4 个子图则共有 40 个子图。

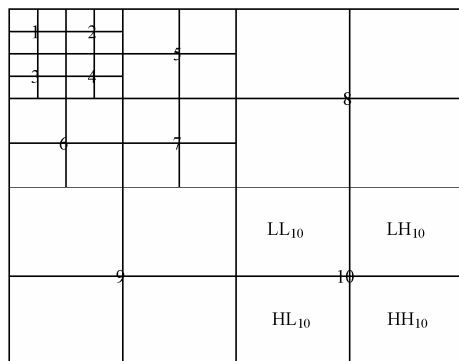


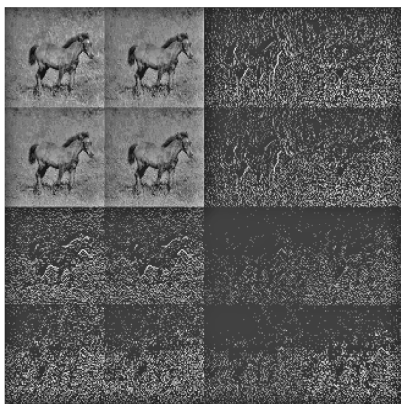
图 4-14 三层多小波分解示意图

2. 多小波直方图

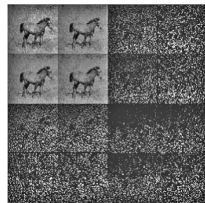
常见的多小波有 GHM 多小波和 CL 多小波，由于 GHM 多小波比其他多小波分解均匀，尚赵伟^[120]提出了基于 GHM 多小波直方图检索算法，并且比较了其他多小波的检索性能。在进行多小波分解前，需要对原始图像进行预处理，常见的多小波预处理方法有重复行法和临界法，其中采用重复行法预处理的多小波变换子带大小与原图相同，而采用临界法预处理的多小波子带大小为原图的 1/2。图 4-15 所示为草原马示例图像，图 4-16 所示为 GHM 多小波的两两种预处理单层分解的示例图像。



图 4-15 草原马示例图像



(a) 重复行预处理GHM单层分解



(b) 临界法预处理GHMAP单层分解

图 4-16 GHM 多小波的两种预处理单层分解的示例图像

由图 4-16 可知,采用重复行法预处理的 GHM 多小波变换子带含有丰富的纹理信息,且每个子带的大小与原图相同。下面将采用重复行法预处理的 GHM 多小波提取多小波纹理直方图特征。首先对图像进行分解,可以得到不同尺度和不同方向的分解信息,低频信息反映了纹理的基本信息,中、高频信息反映了纹理的大部分重要的细节信息。首先对每个小波系数矩阵求绝对值得到小波能量矩阵,再对小波能量矩阵进行组合。尚赵伟等提出了以下 4 种方案。

- (1) 1, 2, 3, 4, 5, 6, 7, 8, 9, 10;
- (2) 1, 3+6, 4+7, 2+5, 8, 9, 10;
- (3) 1, 2+3+4, 5+6+7, 8, 9, 10;
- (4) 1, 2+3+4, 5+6+7, 2+5, 4+7, 3+6, 8, 9, 10。

其次进行 0/1 量化。经过组合后的每个子带仍然含有 4 个子图,以子图为单位对各小波能量矩阵进行 0/1 量化,得到了 0/1 量化矩阵,即

$$F(i, j) = \begin{cases} 1 & \text{当 } |W(i, j)| \geq T \\ 0 & \text{其他} \end{cases} \quad (4-54)$$

式中, T 为阈值,可以采用各子图的均值作为阈值 T ,使阈值具有一定的自适应能力。

最后,按照图 4-17 所示的方法生成数据集,用于表示相应点在不



同量化矩阵的取值。以方案（2）为例，量化矩阵 A_7, A_6, \dots, A_1 （见图 4-2）在相同位置点的值（0/1）组成一个 7 位二进制数，转换为十进制数的取值范围为 0~127，这样就得到 16384 个数据。

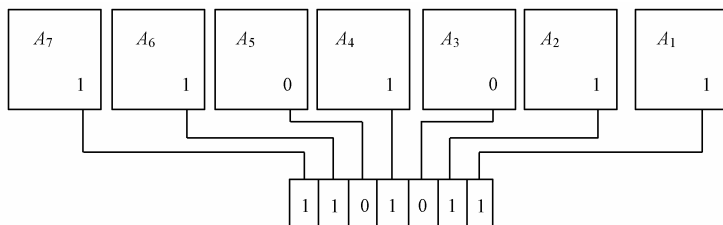


图 4-17 多小波直方图生成数据集

实验分析表明，基于重复行预处理方法得到的多小波直方图算法具有较高的检索性能，比传统小波直方图算法的检索率提高了 16%。然而，该算法没有考虑到图像纹理特征的空间分布特性。

4.3.7 双正交小波变换纹理分析

双正交小波因其所具有的良好特性（如线性相位、高阶消失矩等）被广泛地应用于图像压缩领域。魏海^[121]提出了基于反对称双正交小波的图像检索算法，并且提出了方向梯度相角直方图统计特征量用于表征图像的内容。理论分析和实验结果表明，该算法不仅具有较高的检索效率，同时具有强的光照变化鲁棒性和一定程度的抗几何变化（尺度、平移和旋转等）的能力。

根据 Mallat 的讨论，用于边缘提取的小波函数必须为 $\psi(x)$ 反（奇）对称，它对应的尺度函数 $\phi(x)$ 为偶对称。由小波理论可知，满足这些条件的紧支撑集正交小波仅有 Haar 小波，正交小波可看做是双正交小波的特例。令 $\psi(x)$ 是一个具有反对称性的双正交小波函数， $\theta(x)$ 是 $\psi(x)$ 的积分，即

$$\theta(x) = \int_{-\infty}^x \psi(t) dt \quad (4-55)$$

两边进行傅里叶变换。利用二尺度方程的频域表达式可得^[122]，



$$\hat{\theta}(\omega) = \hat{\psi}(\omega) / (j\omega) = G(\omega/2) \hat{\phi}(\omega/2) / (j\omega) \quad (4-56)$$

其中, $\hat{\psi}(\omega)$ 是和 小波函数 $\psi(x)$ 相对应的尺度函数 $\phi(x)$ 的 复式变换, $\phi(x)$ 满足

$$\psi(x) = 2^{1/2} \sum_k g(k) \phi(2x - k) \quad (4-57)$$

$G(\omega)$ 定义为 $G(\omega) = 2^{1/2} \sum_k g(k) \cdot e^{-j\omega k}$ 。对于一个具备线性相位的反对称双正交小波而言, 系数 $g(k)$ 必须是偶数长度的反对称序列。假设其对称点为 -0.5 , 则

$$g(k) = -g(-1 - k) \quad k = 0, 1, \dots, n-1 \quad (4-58)$$

根据式 (4-58), $G(\omega)$ 可写为

$$\begin{aligned} G(\omega) &= 2^{-1/2} \sum_{k=0}^{n-1} g(k) \cdot \left| e^{-j\omega k} - e^{-j\omega(-1-k)} \right| \\ &= 2^{-1/2} j e^{j\omega/2} \sum_{k=0}^{n-1} g(k) \cdot \left| e^{-j\omega(k-0.5)} - e^{j\omega(k+0.5)} \right| \\ &= 2^{-1/2} j e^{j\omega/2} \sum_{k=0}^{n-1} g(k) \cdot \sin(k + 0.5) \omega \\ &= 2^{-1/2} j \omega e^{j\omega/2} \sum_{k=0}^{n-1} g(k) \cdot (k + 0.5) \end{aligned} \quad (4-59)$$

将式 (4-59) 代入式 (4-56), 可得,

$$\hat{\theta}(\omega) \sim \eta e^{j\omega/4} \hat{\phi}(\omega/2) \quad (4-60)$$

式中, $\eta = 2^{1/2} \sum_{k=0}^{n-1} g(k) \cdot (k + 0.5)$ 。由此得到

$$\theta(x) \sim \eta \phi(2x + 0.5) \quad (4-61)$$

式 (4-61) 表明, 在一阶近似下反对称双正交小波函数 $\psi(x)$ 的积分 $\theta(x)$ 和尺度函数 $\phi(2x)$ 比例关系。

采用一定的反对称双正交小波对图像进行塔式分解, 并通过半重构和局部模极大值检测过程可得到多个分辨率级上的边缘图像, 每个分辨率的边缘图像中的每个边缘像素都具有一个梯度模值和一个方向梯度相角值, 假设第 j 级分辨率上的每一边缘像素点所具有模值为



$$M_j(x, y) = \left[\left(\frac{\partial f_j(x, y)}{\partial x} \right)^2 + \left(\frac{\partial f_j(x, y)}{\partial y} \right)^2 \right]^{1/2} \quad (4-62)$$

方向梯度相角为

$$A_j(x, y) = \arctan \left[\left(\frac{\partial f_j(x, y)}{\partial y} \right) / \left(\frac{\partial f_j(x, y)}{\partial x} \right) \right] \quad (4-63)$$

式中, $f_j(x, y)$ 为图像的第 j 级近似, x 和 y 表示了边缘像素点在模图中的位置。

利用每一分辨率级上的边缘图像可计算方向梯度相角直方图。首先, 为了体现图像中的“强边缘”(图像中大的目标、景物的轮廓和边界)并尽量避免噪声和精细纹理所造成的“弱边缘”的影响, 引入比率变量 R , 定义为用于方向梯度相角直方图统计的边缘像素的个数和该分辨率级上的模图像素点总数的比率, 用来控制参与方向梯度相角直方图统计计算的边缘像素个数; 实际计算直方图时, 首先预设一定的 R 值, 从 7 级分辨率上的边缘像素集合中选出和 R 比率相对应个数的模值最大的边缘像素子集, 将该边缘像素子集按其方向梯度相角分布分成 C 类(C 是预设的类别数, 用于控制直方图分量的个数), 并统计每类中边缘像素的个数, 最终得到具有 C 个分量的方向梯度相角直方图特征矢量。图像的方向梯度相角直方图还具有许多良好的特性, 对光照变化的鲁棒性、平移、尺度不变性。

4.3.8 复数小波变换纹理分析

为了克服小波变换在提取图像纹理特征时存在振荡、平移变化、混频和缺乏方向性等问题, 张德胜^[123]提出了基于双树复小波变换的图像检索算法。为了降低离散小波变换的平移敏感性, Kingsbury 等^[124]提出了二维双树复小波变换 (Dual Tree Complex Wavelet Transform, DTCWT)。二维双树复小波变换包含两个平行的小波树, 即树 A 和树 B 两个分支。其中上部树 A 的叠加滤波器组表示复小波变换的实部, 下



部树 B 的叠加滤波器组表示复数小波变换的虚部。图 4-18 所示的是一维双树复小波变换的示意图。

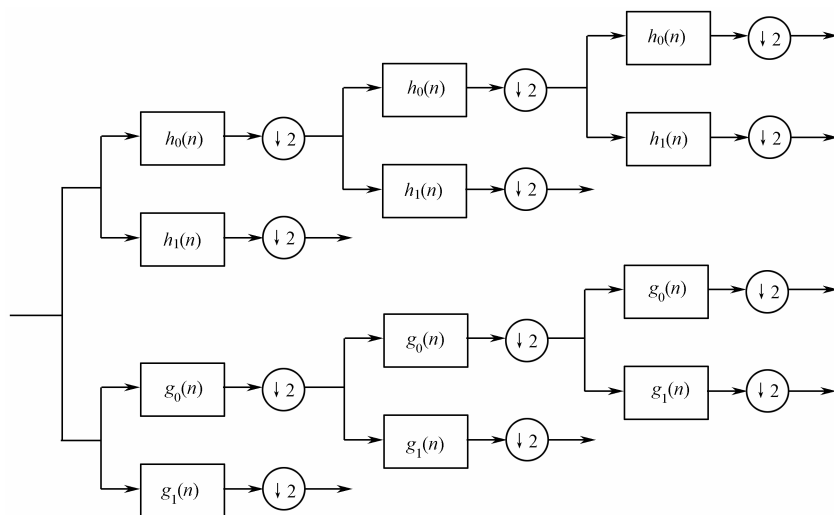


图 4-18 一维双树复小波变换

二维双树复小波变换可产生 6 个方向的高频子图像, 6 个小波子带的主方向分别为 $\theta = \{+15, +45, +75, -15, -45, -75\}$, 这样就可以提取图像在 6 个方向上的细节系数矩阵, 实现纹理特征的分析。图 4-19 所示为双树复小波变换针对 Zoneplate 图像的 3 层分解示意图, 其中 4.19(b) ~ (d) 每层分解的上 3 幅图像是复数小波变换的实部, 而下 3 幅图像是虚部。

利用双树复小波变换的灰度共生矩阵方法提取纹理特征可分为如下 3 个步骤。

- (1) 将图像转换成 256 色灰度图;
- (2) 对查询图进行双树复小波变换;

(3) 计算查询图的灰度共生矩阵并求其统计量特征, 再对这些特征求取均值和方差, 归一化并保存数据到数据库。

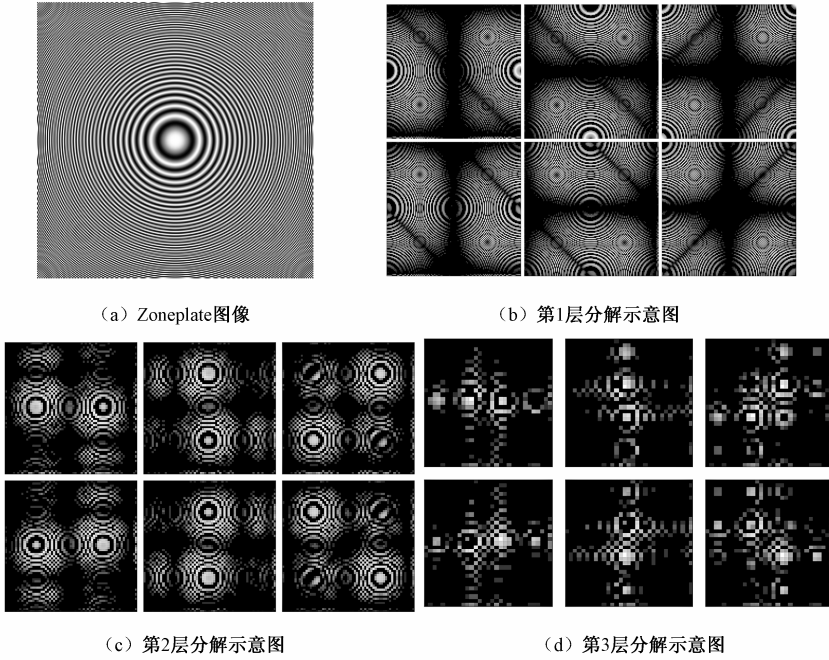


图 4-19 双树复小波变换分解示意图

可以提取灰度共生矩阵的 4 个的统计量，即二阶矩（ASM）、对比度（Con）、相关度（Cor）和熵（Ent），其计算公式为

$$ASM = \sum_{i=1}^{N_q} \sum_{j=1}^{N_q} \{P_{ij}\}^2 \quad (4-64)$$

$$Con = \sum_{n=0}^{N_q-1} n^2 \left\{ \sum_{i=1}^{N_q} \sum_{j=1}^{N_q} P_{ij} \right\}, |i-j|=n \quad (4-65)$$

$$Cor = \frac{\sum_{i=1}^{N_q} \sum_{j=1}^{N_q} i \cdot j \cdot P_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (4-66)$$

$$Ent = - \sum_{i=1}^{N_q} \sum_{j=1}^{N_q} P_{ij} \log_2 [P_{ij}] \quad (4-67)$$

式中， N_q 为量化的灰度级， P_{ij} 为灰度共生矩阵的系数， μ_x 和 σ_x 是 P_{ij} 中每一行数据的均值和方差， μ_y 和 σ_y 是 P_{ij} 中每一列数据的均值和方差。



实验分析表明，双树复小波结合灰度共生矩阵法具有检索的有效性，然而其检索算法具有较大的计算复杂度。

4.3.9 双密度小波变换纹理分析

为了克服传统小波变换的时移敏感性，Selesnick^[125]提出了双密度小波变换（Double Density Wavelet Transform），其双尺度方程可表示为

$$\phi(t) = \sqrt{2} \sum_n h_0(n) \phi(2t - n) \quad (4-68)$$

$$\psi_i(t) = \sqrt{2} \sum_n h_i(n) \phi(2t - n) \quad i = 1, 2 \quad (4-69)$$

式中， $h_0(n)$ 是低频滤波器， $h_i(n)(i=1,2)$ 是高频滤波器。双密度小波变换采用了过采样方案，具有近似时移不变性的优点，同时采用了两个高频滤波器和一个低频滤波器，与传统小波变换相比，具有较多的方向信息，已被应用于图像处理领域。双密度小波针对2D信号的分解采用可分离滤波器对行和列分别滤波^[125-126]，这样针对纹理图像的单层分解可以得到9个方向子带。图4-20所示为辣椒图像，图4-21所示为针对图4-20的单层双密度小波分解示意图。



图 4-20 辣椒图像

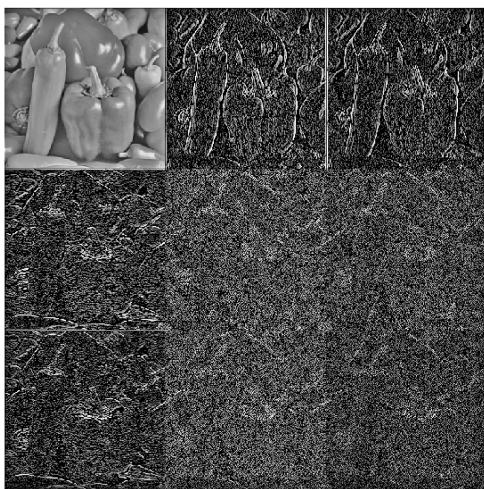


图 4-21 辣椒图像的单层双密度小波分解



双密度小波变换具有良好的方向性和平移不变性，尚赵伟等^[126]提出了基于双密度小波变换的纹理图像检索算法。由于双密度小波变换分解的子带信息较多，可以将双密度小波变换的子带进行系数组合，然后提取组合后的子带纹理特征，从而进行图像检索。由于在双密度小波变换中，子带 $h_y^2 h_x^2$ 、 $h_y^2 h_x^1$ 、 $h_y^1 h_x^2$ 和 $h_y^1 h_x^1$ 与传统小波分解变换的 HH 子带具有相同的频域支撑，其半延迟特性隐含着相应 $h_y^2 h_x^1$ 子带列的小波系数的空域支撑，位于 $h_y^2 h_x^2$ 子带列的半整数的右边，因此可以将子带 $h_y^2 h_x^2$ 、 $h_y^2 h_x^1$ 、 $h_y^1 h_x^2$ 和 $h_y^1 h_x^1$ 的小波系数根据以下公式进行组合。

$$\left. \begin{aligned} h_y^2 h_x^2 \left(\frac{m}{2}, \frac{n}{2} \right) & \quad m, n = 2, 4, 6, \dots \\ h_y^2 h_x^1 \left(\frac{m}{2}, \frac{n-1}{2} \right) & \quad m = 2, 4, 6, \dots, n = 1, 3, 5, \dots \\ h_y^1 h_x^2 \left(\frac{m-1}{2}, \frac{n}{2} \right) & \quad m = 1, 3, 5, \dots, n = 2, 4, 6, \dots \\ h_y^1 h_x^1 \left(\frac{m-1}{2}, \frac{n-1}{2} \right) & \quad m, n = 1, 3, 5, \dots \end{aligned} \right\} \quad (4-70)$$

利用式 (4-70) 形成了新的子带 HH_{DD} ，该子带具有与传统小波变换 HH 子带相同的频域支撑，但其大小是 HH 子带的 4 倍。同理，可以根据以下公式进行组合。

$$\left. \begin{aligned} h_y^0 h_x^2 \left(m, \frac{n}{2} \right) & \quad n = 2, 4, 6, \dots \\ h_y^0 h_x^1 \left(m, \frac{n-1}{2} \right) & \quad n = 1, 3, 5, \dots \\ h_y^2 h_x^0 \left(\frac{m}{2}, \frac{n}{2} \right) & \quad m = 2, 4, 6, \dots \\ h_y^1 h_x^0 \left(\frac{m}{2}, \frac{n-1}{2} \right) & \quad m = 1, 3, 5, \dots \end{aligned} \right\} \quad (4-71)$$

这样就生成了子带 HL_{DD} (LH_{DD})，该子带也具有与传统小波变换 HL



(LH) 子带相同的频域支撑,但其大小是 HL (LH) 子带的 2 倍。只有子带 $h_y^0 h_x^0$ 的频域支撑与传统小波变换 LL 子带相同且大小一致,才可以将 $h_y^0 h_x^0$ 表示成 LL_{DD} 。这样,经过组合后的双密度小波变换就可以计算图像的纹理特征,同时其特征数量与传统小波变换相同。

文献[126]提出利用双密度小波的子带统计矩和共生矩阵作为图像的纹理特征。实验分析表明,组合后的双密度小波变换同时利用一阶和二阶统计矩进行纹理检索具有较好的检索性能,其性能优于传统的小波变换。目前,双密度小波变换已成功应用于图像去噪、图像融合等方面^[127-128],基于双密度小波变换的图像处理应用是今后研究的热点问题。

4.4 旋转不变性纹理特征

4.4.1 Polar-wavelet 纹理分析

Pun^[129]采用图像的极坐标变换和平移不变小波变换提取了具有旋转不变的纹理特征,并将其应用于基于纹理特征的图像检索。

Polar-wavelet 图像检索算法的流程图如图 4-22 所示。

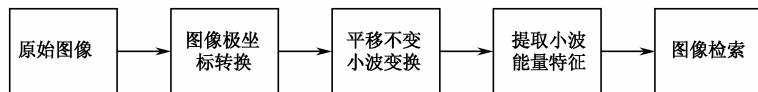


图 4-22 Polar-wavelet 图像检索算法的流程图

常见的平移不变小波变换是平稳离散小波变换 (Stationary Wavelet Transform, SWT), 也称为静态小波, 其一维信号分解的平稳离散小波变换流程如图 4-23 所示。从图 4-23 中可以看到, 在平稳小波变换过程中不存在下采样的问题, 因此平稳小波每次分解的系数长度与原始信号的长度相同, 能够更好地在空间上与原始信号相对应, 有利于处理具有统计规律的信号。对于图像处理普遍采用的是二维离散小波变换, 二维小波变换是一维小波变换的推广。图 4-24 所示为草原马图像的二维平稳离散小波变换 3 层分解示意图。

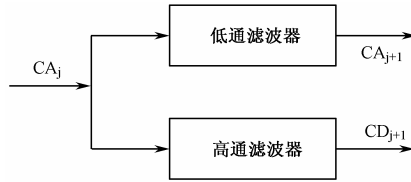


图 4-23 平稳离散小波变换的单步分解

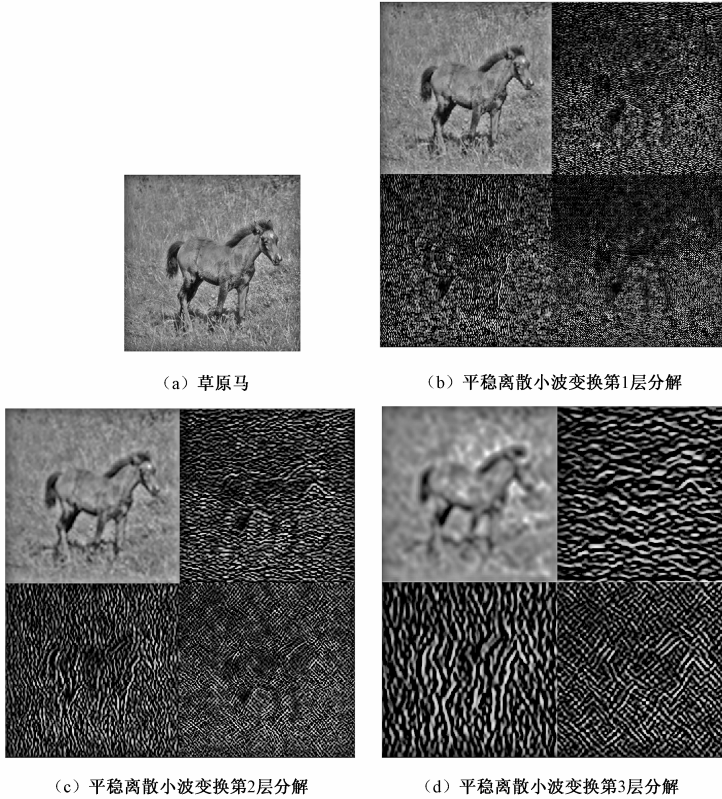


图 4-24 平稳离散小波变换 3 层分解示意图

图像极坐标变换是指将图像从笛卡儿坐标系转换至极坐标系，这样，笛卡儿坐标系下的图像匹配就转化为极坐标下的图像匹配问题。图像 $f(x, y)$ 到 $f(r, \theta)$ 的极坐标变换定义见式 (4-72)，其中 (x_c, y_c) 为变换中心。

$$\begin{cases} r = \sqrt{(x - x_c)^2 + (y - y_c)^2} \\ \theta = \arctan \frac{y - y_c}{x - x_c} \end{cases} \quad (4-72)$$



由式(4-72)可知一幅图像 $f(x,y)$ (大小为 $N \times N$) 的极坐标可以表示为 $p(r,\theta)$, 同时图像的旋转就相当于极坐标 θ 的平移。文献[129]使用的图像极坐标变换为

$$p(i,j) = f\left(\left[\frac{N}{2}\right] + \left[j \cos\left(\frac{2\pi i}{S}\right)\right], \left[\frac{N}{2}\right] - \left[j \sin\left(\frac{2\pi i}{S}\right)\right]\right) \quad (4-73)$$

式中, $i=0,1,2,\dots,S-1$, $j=0,1,2,\dots,\left[\frac{N}{2}\right]-1$, S 为沿径向的采样频率。

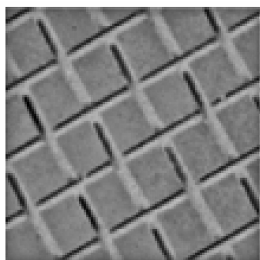
因此, 采用平移不变小波对图像 $p(r,\theta)$ 进行变换, 就得到了图像的旋转不变的小波分解系数, 进一步可以提取图像的小波能量特征。设子带 j 的图像为 $I(x,y)$, 则子带 j 下的能量值 E_j 可定义为

$$E_j = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N I^2(x,y) \quad (4-74)$$

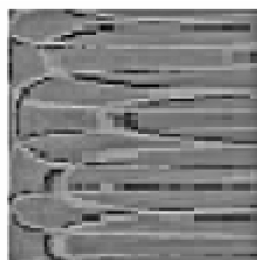
实验结果表明, Polar-wavelet 纹理检索算法具有旋转不变性, 但仍然不具有尺度不变性, 因此在特征提取前需要将原始图像进行尺度归一化预处理。

4.4.2 Log-polar变换纹理分析

为了进一步获得具有旋转和尺度不变的图像纹理特征, Pun^[130]等人提出采用对数极坐标变换(Log-polar)进行了纹理分析。图像对数极坐标变换是指将图像从笛卡儿坐标系转换至对数极坐标系, 这样, 笛卡儿坐标系下的图像匹配就转化为对数极坐标下的图像匹配问题。图 4-25 所示为纹理图像的 Log-polar 变换。



(a) 纹理示例图像



(b) Log-polar图像

图 4-25 纹理图像的 Log-polar 变换



图像 $f(x, y)$ 到的 $\text{lp}(r, \theta)$ 对数极坐标变换定义见式 (4-75)，其中 (x_c, y_c) 为变换中心。

$$\begin{cases} r = \ln \sqrt{(x - x_c)^2 + (y - y_c)^2} \\ \theta = \arctan \frac{y - y_c}{x - x_c} \end{cases} \quad (4-75)$$

图像 $f(x, y)$ (大小为 $N \times N$) 的对数极坐标表示可以表示为 $\text{lp}(r, \theta)$ ，图像的旋转就相当于图像极坐标 θ 的平移，图像尺度变换相当于 r 方向的平移。由文献[130]可知，对于一幅 $S \times R$ 的 Log-polar 图像 (原始图像大小为 $N \times N$) 的对数极坐标变换也可以表示为

$$\text{lp}(i, j) = p(i, [\frac{\log_2(j+2)}{\log_2(R+2)} \cdot [\frac{N}{2}]]]) \quad (4-76)$$

式中， $i=0,1,2,\dots,S-1$ ， $j=0,1,2,\dots,R-1$ ， S 为沿径向的采样频率。这样，沿 r 方向 $p(r, \theta)$ 进行平移不变小波变换，就可以得到平移不变小波变分解，而沿 θ 方向对 $p(r, \theta)$ 进行平移不变小波分解就可以得到旋转不变的小波分解。文献[130]采用了 5 种能量特征的统计值来表征图像的纹理，设 C_k 为小波分解系数， μ 表示小波分解系数的均值，能量公式见表 4-1。

表 4.1 五种小波统计特征

Norm-2 能量	$e_1 = \frac{1}{N^2} \sum_{k=1}^N C_k ^2$
Norm-1 能量	$e_2 = \frac{1}{N^2} \sum_{k=1}^N C_k $
标准差	$e_3 = \sqrt{\frac{1}{N^2} \sum_{k=1}^N (C_k - \mu)^2}$
平均残差	$e_4 = \sum_{k=1}^N C_k - \mu $
能量熵	$e_5 = -\frac{1}{N^2} \sum_{k=1}^N C_k ^2 \log C_k ^2$

实验结果表明，Log-wavelet 纹理检索算法具有尺度和旋转不变性，针对几何变换图像库，具有较好的检索性能，但针对自然图像的检索性能较差。

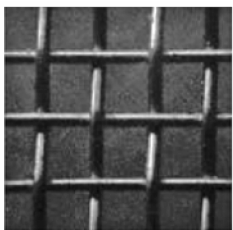


4.4.3 RIM纹理分析算法

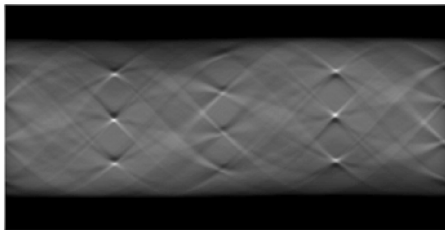
为了寻求具有旋转不变的纹理特征，Kourosh^[131]提出了基于 Radon 和平移不变小波的纹理分析算法（RIM 算法），对高斯噪声具有较强的鲁棒性。二维函数 $f(x, y)$ 的 Radon 变换可以表示为该函数沿包含该函数的平面的一族直线的线积分，即

$$R\{f(x, y)\} = \iint f(x, y) \delta(t - x \cos \theta - y \sin \theta) dx dy = g_{\theta}(t) \quad (4-77)$$

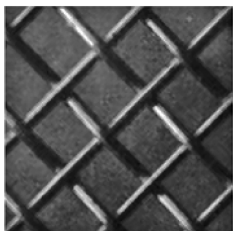
式中， $t = x \cos \theta + y \sin \theta$ ，沿着一系列平行线的积分就组成了投影 $g_{\theta}(t)$ ，投影集合 $\{g_{\theta}(t), \theta \in [0, \pi]\}$ 即为 Radon 变换。



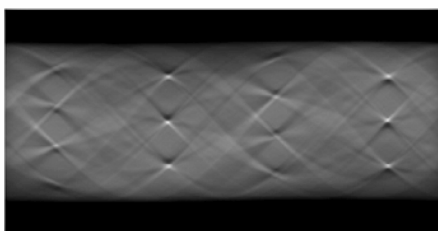
(a) 纹理图像



(b) 图 (a) 的 Radon 变换



(c) 图 (a) 的旋转图像



(d) 图 (c) 的 Radon 变换

图 4-26 不同旋转图像的 Radon 变换

图 4-26 所示为不同旋转图像的 Radon 变换，可知原图像的旋转就等于 θ 角度的循环平移。Kourosh^[131]等人针对图像 Radon 变换的投影数据，针对 θ 方向采用平移不变小波变换进行分解，而针对 t 方向仍然可以采用普通小波变换进行分解，就可以得到旋转不变的小波分解。采用图像小波分解的能量特征表示纹理性，因此可以得到具有旋转不变的纹



理特征。设子带大小为 $M \times N$ ，图像子带 j 的小波能量特征可以表示为

$$E_j^1 = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N |I(x, y)| \quad (4-78)$$

$$E_j^2 = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N I^2(x, y) \quad (4-79)$$

然而，RIM 算法不具有平移和尺度不变性。因此，文献[132]在平移和尺度不变小波变换（Translation- and Scale-Invariant wavelet, TSI wavelet）基础上提出了改进的 RIM 算法。

TSI wavelet 变换^[133]是根据几何矩原理构造的平移和尺度不变的小波变换。对信号 $f(x)$ 而言，若 $f(x)$ 的多尺度小波分解系数和 $f(ax+b)(a>0, b \in R)$ 的分解系数相同，则小波分解是平移和尺度不变的。由于实际信号都是有限的，则信号 $f(x)$ 的一阶矩和二阶矩均存在^[133-134]。

$$m_f = \int_{-\infty}^{+\infty} f(x)dx, \mu_f = \frac{1}{m_f} \int_{-\infty}^{+\infty} xf(x)dx \quad (4-80)$$

$$\sigma_f^2 = \frac{1}{m_f} \int_{-\infty}^{+\infty} (x - \mu_f)^2 f(x)dx \quad (4-81)$$

$$\text{设 } \phi_{j,k}^f(x) = \frac{2^{-j/2}}{\sqrt{\sigma_f}} \phi\left(2^{-j} \frac{x - \mu_f}{\sigma_f}\right), \psi_{j,k}^f(x) = \frac{2^{-j/2}}{\sqrt{\sigma_f}} \psi\left(2^{-j} \frac{x - \mu_f}{\sigma_f}\right) (j, k \in Z)$$

生成的函数空间为 $\bar{V}_f^f = \text{span}\{\bar{\phi}_{j,k}^f(x), j, k \in Z\}$, $\bar{W}_f^f = \text{span}\{\bar{\psi}_{j,k}^f(x), j, k \in Z\}$, 由 $\psi_{j,k}(x)$ 和 $\phi_{j,k}(x)$ 的正交规范性，可得 $\bar{\phi}_{j,k}^f(x)$ 和 $\bar{\psi}_{j,k}^f(x)$ 分别是函数空间 \bar{V}_f^f 和 \bar{W}_f^f 的正交规范基。 $\{\bar{V}_f^f\}_{j \in Z}$ 构成 $L^2(R)$ 的一个多尺度逼近，且 $\bar{V}_{j-1}^f = \bar{V}_f^f \oplus \bar{W}_f^f$ 。信号 $f(x)$ 在 \bar{V}_f^f 和 \bar{W}_f^f 上的小波分解为

$$f(x) = \sum_{k=-\infty}^{+\infty} \overline{c_{J,K}^f} \bar{\phi}_{J,K}^f(x) + \sum_{j=1}^J \sum_{k=-\infty}^{+\infty} \overline{d_{j,k}^f} \bar{\psi}_{j,k}^f(x) \quad (4-82)$$

式中， $\overline{c_{J,K}^f} = \langle f, \bar{\phi}_{J,K}^f \rangle = \sqrt{\sigma_f} \int_{-\infty}^{+\infty} f(\sigma_f x + \mu_f) \phi_{J,K}(x) dx$, $\overline{d_{j,k}^f} = \langle f, \bar{\psi}_{j,k}^f \rangle = \sqrt{\sigma_f} \int_{-\infty}^{+\infty} f(\sigma_f x + \mu_f) \psi_{j,k}(x) dx$ 。



【定理 1】 设 $f(x) \geq 0, \int_{-\infty}^{+\infty} f(x)dx < +\infty$ 且 $g(x) = f(ax+b)$ ($a \geq 0, b \in R$), 则

$$\frac{\overline{c_{j,k}^f}}{\sigma_f^{\frac{1}{2}}} = \int_{-\infty}^{+\infty} f(\sigma_f x + \mu_f) \phi_{j,k}(x) dx = \int_{-\infty}^{+\infty} g(\sigma_g x + \mu_g) \phi_{j,k}(x) dx = \frac{\overline{c_{j,k}^g}}{\sigma_g^{\frac{1}{2}}} \quad (4-83)$$

$$\frac{\overline{d_{j,k}^f}}{\sigma_f^{\frac{1}{2}}} = \int_{-\infty}^{+\infty} f(\sigma_f x + \mu_f) \psi_{j,k}(x) dx = \int_{-\infty}^{+\infty} g(\sigma_g x + \mu_g) \psi_{j,k}(x) dx = \frac{\overline{d_{j,k}^g}}{\sigma_g^{\frac{1}{2}}} \quad (4-84)$$

由上述定理可以得到非负信号 $f(x)$ 重整后的标准信号 $f(\sigma x + \mu)$, 且 $f(\sigma x + \mu)$ 在小波多尺度空间 V_j 和 W_j 的小波变换系数具有平移和尺度不变性, 可以认为是 $f(x)$ 的平移、尺度不变小波分解。平移和尺度不变小波变换的实质是针对信号进行自适应重整, 得到“标准信号”后再进行小波变换就可以获得平移和尺度不变的小波分解。

RIM 方法只针对 Radon 变换投影矩阵的 θ 方向进行了平移不变小波, 而在 t 方向使用了普通小波, 因此仅具有旋转不变性。由 Radon 变换的平移性质可知, 原图像的平移就等于 t 的平移, 可以采用文献[97]的方法, 利用投影质心对投影矩阵进行中心化以消除平移影响。当图像发生尺度变换时, Radon 投影变化如第三章式 (3-65) 所示, 不能够直接采用 TSI 小波分解。

当图像产生尺度变换时, Radon 投影矩阵的奇异值具有尺度相关性, 因此可以利用图像的奇异值对投影数据规范化。设 λ 表示比例因子, 矩阵的奇异值为 η , 当图像产生尺度变换时, Radon 投影矩阵奇异值变化为 $\eta_k^s = \frac{\eta_k}{\lambda^{3/2}}$ 。

令 $h(t) = \frac{g(t)}{[\max(\eta_i)]^{2/3}} (i = 0, 1, \dots, k)$, 则

$$h^s(t) = \frac{g^s(t)}{[\max(\eta_i^s)]^{2/3}} = \frac{\frac{1}{\lambda} g_\theta(\lambda t)}{\left[\max \left(\frac{\eta_i}{\lambda^{3/2}} \right) \right]^{2/3}} = \frac{g_\theta(\lambda t)}{[\max(\eta_i)]^{2/3}} = h(\lambda t) \quad (4-85)$$



因此, $h(t)$ 满足自适应 TSI 小波分解条件。改进的 RIM 算法针对 Radon 投影矩阵的 θ 方向进行平移不变小波, 可以获得旋转不变的小波分解; 而针对 t 方向使用平移和尺度不变的小波分解, 可以获得平移和尺度不变的小波分解, 这样图像就得到了具有旋转、平移和尺度不变的小波分解, 最后利用各子带系数能量的均值和方差作为图像纹理特征。

实验分析表明, 改进的 RIM 算法具有旋转、平移和尺度不变性, 针对几何变换图像库具有较好的检索性能, 但针对自然图像而言, 检索性能不够理想。

4.5 本章小结

纹理是一种不依赖于图像颜色和亮度的反映图像同质现象的视觉特征, 体现了物体表面的内在特征。本章分析了空间域的纹理分析和频率域的纹理分析方法, 重点介绍了频率域的纹理分析方法。空间域的纹理分析方法包括 Tamura 纹理特征、灰度共生矩阵法、LBP 算法和 Laws 纹理能量法等。频率域的纹理分析方法主要包括傅里叶功率谱法、小波变换统计矩、小波直方图、小波 GGD&KLD 方法、Gabor 小波纹理分析、多小波变换纹理分析、双正交小波变换纹理分析、复数小波变换和双密度小波变换纹理分析方法。同时, 分析了 3 种基于频域变换的旋转不变纹理分析方法, 包括 Polar-wavelet 纹理分析、Log-polar 变换纹理分析和 RIM 纹理分析算法。目前, 基于频率域的纹理分析方法已经成为当前图像纹理分析的研究热点。

第五章



基于感兴趣区域的图像检索

5.1 感兴趣区域检索概述

感兴趣区域是指图像中最能引起人们兴趣、最能表现图像内容的区域。基于感兴趣区域的图像检索方法在一定程度上实现了物体层次的图像检索，消除了图像的冗余信息，突出了图像的主要内容，从而在一定程度上克服了语义鸿沟，是目前基于内容图像检索研究的热点之一。典型的基于区域的图像检索系统有 NeTra 系统、Blobworld 系统及 SIMPLIcity 系统等。基于感兴趣区域的图像检索方法首先要提取感兴趣物体或目标的区域，然后再利用视觉感知特征进行检索，因此算法往往受到图像分割技术限制。本章将介绍 3 种基于感兴趣区域的图像检索算法，其中包括基于视觉显著度模型的检索算法、基于小波感兴趣区域图像检索，以及基于颜色和熵的感兴趣区域的检索算法。

5.2 基于视觉显著度模型的图像检索

为了提取图像的感兴趣区域，Itti^[135]提出了一个有效的视觉注意计算模型，该模型在不同尺度上分别计算 3 个特征通道（颜色、亮度和方向）的注意程度定量值，然后线性组合它们，通过迭代过程形成最终的显著度图，显著度图记录了图像中每一个点受关注的程度，亮度越大的



点意味着受关注的程度也越大，如图 5-1 所示。

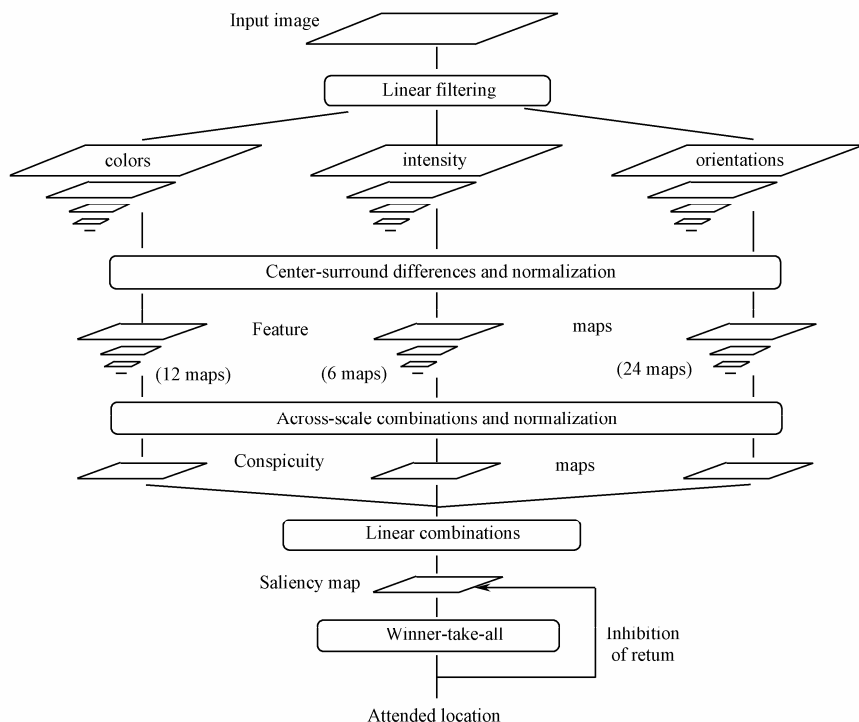


图 5-1 Itti 视觉注意度模型

Itti 视觉注意度模型的工作原理如下所述。

- (1) 将输入图像进行线性滤波。
- (2) 提取不同尺度的颜色、亮度和方向特征。
- (3) 将各维特征为统一尺度多个级别的特征图。

(4) 计算中央周边差，归一化得到 3 个特征的视觉关注度图，将视觉关注度图进行线性组合为视觉显著度图，视觉显著度图利用 WTA 竞争机制，找到注意焦点。

图 5-2 所示的是示例图像的视觉显著度图，图中点的亮度越大意味着它受关注程度越强。

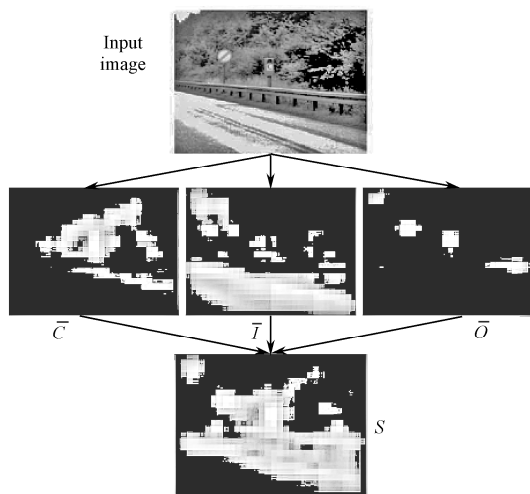


图 5-2 图像的视觉显著度图

韩军伟等^[98]利用图像的视觉显著度图,结合边缘特征和种子区域增长算法提取了图像的感兴趣区域,并利用感兴趣区域的颜色和纹理特征进行检索。假设图像中的一点 (x,y) ,根据视觉注意计算模型得出的显著度为 $s(x,y)$,本算法中使用的关注度由式(5-1)来计算:

$$\hat{s}(x,y) = s(x,y) \otimes \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (5-1)$$

文献[98]结合视觉显著度图和种子区域增长算法来自动提取图像中感兴趣的物体,本算法共需5个步骤。

1. 颜色量化

图像中的颜色被量化成若干能够区分不同区域的代表颜色类。首先采用非线性滤波器^[136]进行滤波,滤波后的值能够指示出局部区域的平滑程度;然后再利用修正的Lloyd算法(GLA)^[137]来量化像素的颜色;最后利用一个聚类算法(Agglomerative Clustering Algorithm)来合并相似的颜色聚类,最终使得聚类中心之间的最小距离小于某一个预先设置的阈值。

颜色量化后,每一个颜色量化类赋予一个标志。图像中每一个像素



的颜色用它对应的标志代替后就形成了一个颜色类图像（Color Class Map）。

2. 相对位置指示图

在颜色类图像的基础上生成一个相对位置指示图，它能够估测出每一个像素点的相对位置是处于某一区域的内部还是更靠近边缘。设 Z 表示颜色类图像中的所有 N 个像素点集合， $z = (x, y), z \in Z$ ， m 为均值，

$$m = \frac{1}{n} \sum_{z \in Z} z \quad (5-2)$$

假定集合 Z 被分为 C 个颜色类 $Z_i, i=1, 2, \dots, C$ 。假定颜色类 Z_i 中点的个数为 N_i ，则平均值 m_i 可计算为

$$m_i = \frac{1}{n} \sum_{z \in Z_i} z \quad (5-3)$$

定义 S_w 表示所有类的方差总和，即

$$S_T = \sum_{z \in Z} \|z - m\|^2 \quad (5-4)$$

$$S_w = \sum_{i=1}^C S_i = \sum_{i=1}^C \sum_{z \in Z_i} \|z - m_i\|^2 \quad (5-5)$$

式中，根据 S_T 和 S_w ，可以定义相对位置指示 RPD 为

$$RPD = (S_T - S_w) / S_w \quad (5-6)$$

若图像区域的不同颜色类之间的距离较大，则 RPD 就会较大，而颜色类在整幅图像中分布较为平均，则 RPD 的取值就会比较小。因此，对图像中一个像素点而言，以其为中心选取一个 $l \times l$ 大小的小窗口，在此小窗口中内计算 RPD 值，这个值的大小可以估计出此像素点的相对位置。如果图像中每一个像素的值都由它的 RPD 值代替，就形成了相对位置指示图，图中亮度越大的点则表示该点越有可能靠近边缘。

3. 感兴趣物体种子区域的选取

种子点应当是物体上直观的和具有代表性的区域，可以根据种子点度量选择最佳种子区域。设 P 是图像中的一个像素点， R 表示以 P 为中



心的 $\lambda \times \lambda$ 大小的区域, RPD 表示相对位置指示, S 表示视觉显著度, μ_{PRD} 表示区域 R 内的 RPD 平均值, σ_{PRD} 表示区域 R 内的 RPD 标准差, μ_s 表示区域 R 内的视觉显著度平均值, σ_s 表示区域 R 内的视觉显著度标准差, 则区域 R 的种子点度量可以定义为

$$A_R = \exp^{-(\mu_{\text{PSD}} + \sigma_{\text{PSD}})} - \exp^{-(\mu_s - \sigma_s)} \quad (5-7)$$

由式 (5-7) 可知, 种子点度量值越大, 则该区域对应的注意度越大, 因此对应的区域就更优先地被选为种子区域。

4. 感兴趣物体生长

从种子区域开始依次增长感兴趣物体, 增长过程由显著度图和相对位置指示图来共同引导和控制。假设一个感兴趣物体 O 从种子区域 R 开始生长, 图像中已经被认定属于感兴趣物体的点称为已标注点, 不属于任何一个物体的点称为未标注点。与物体 O 相邻的点集合可以定义为

$$H = \{(x, y) \in O \mid N(x, y) \cap O \neq \emptyset\} \quad (5-8)$$

式中, $N(x, y)$ 表示以点 (x, y) 为中心的 3×3 的小区域。像素 P 是否归入 O 的相似度可以定义为

$$\text{ST}(P \rightarrow O) = \begin{cases} \text{true} & \text{if } S > T_s \text{ and } \text{PSD} < T_{\text{PSD}} \\ \text{false} & \text{otherwise} \end{cases} \quad (5-9)$$

式中, T_s 和 T_{PSD} 为阈值, 通过测试, P 就可以归入 O 。下面将利用基于信息熵的方法确定最佳阈值, 设感兴趣点和非感兴趣点的概率分布是独立的, 即

$$P_n(t) = \frac{n_t}{\sum_{i=0}^T n_i}, 0 \leq t \leq T \quad (5-10)$$

$$P_a(t) = \frac{n_t}{\sum_{i=T+1}^M n_i}, T+1 \leq t \leq M \quad (5-11)$$

式中, $P_n(t)$ 表示非感兴趣点的概率, $P_a(t)$ 表示感兴趣点的概率。两种类型的点的信息熵可定义为

$$E_n(T) = -\sum_{t=0}^T P_n(t) \log P_n(t) \quad (5-12)$$



$$E_a(T) = - \sum_{t=T+1}^M P_a(t) \log P_a(t) \quad (5-13)$$

能够区分两类阈值 T 的最佳取值 T^* 应该满足^[138]

$$T^* = \arg \max_{T=0,1,\dots,M} \left\{ E_n(T) + E_a(T) \right\} \quad (5-14)$$

这样最佳的阈值 T_s 和 T_{psd} 都可以通过计算得到。

5. 后期处理

实际上, 由于受噪声、显著度图的不准确、相对位置图的不准确、种子点选取的误差等因素的影响, 往往提取出的感兴趣物体并不令人满意, 需要填充物体内部的一些空洞, 可以采用经典的形态学闭运算来完成。

经过上述 5 个步骤就可以提取到图像的感兴趣物体。文献[98]提取了感兴趣的物体的颜色和纹理进行图像检索, 颜色特征采用 RGB 空间的颜色相关图表示, 纹理特征采用 Gabor 小波变换子带的均值和方差表示, 实验结果表明, 针对具有明显感兴趣物体的图像检索具有较高的检索率。

5.3 基于小波感兴趣区域的图像检索

小波变换具有良好的时频分析能力, 已被广泛应用于图像检索中, 王向阳^[139]等提出了基于小波感兴趣区域的图像检索算法。由小波分析理论可知, 在小波变换域内, 感兴趣区的局部小波系数方差要明显大于背景区域的局部小波系数方差。根据该原理, 可以通过定义图像局部区域的小波系数方差来区分感兴趣区和背景, 即提取出感兴趣区。

设原始彩色图像为 $I = \{I(i, j), 0 \leq i \leq N, 0 \leq j \leq M\}$, $I(i, j)$ 代表原始彩色图像的第 i 行、 j 列像素的颜色值。选取适合人眼视觉感知特性的 HSV 颜色空间进行彩色图像的分割与检索, 而且在彩色图像的亮度分量上提取感兴趣区域。感兴趣区提取方法如下所述。



(1) 对彩色图像 I 的亮度分量 V 进行单层小波变换, 得到小波系数矩阵 U 。然后, 将亮度分量 V 划分成互不重叠的图像子块, 令空间域图像子块的大小为 $L \times W$ 。

(2) 在小波系数矩阵 U 的 3 个高频子带中, 计算对应于空间域图像块的 3 个小波系数块的块内系数方差, 并求出其平均值, 以此作为图像分割的特征。设在亮度分量 V 中, 位于 $(i, j), 0 \leq i \leq M/L, 0 \leq j \leq N/W$ 处的空间域图像块为

$$V_{ij} = \{v(mi, nj), 0 \leq m < L, 0 \leq n < W\} \quad (5-15)$$

则 3 个高频子带 LH、LH 和 HH 内相应小波系数块分别为

$$U_{ij}^\lambda = \{U^\lambda(mi, nj), 0 \leq m < l/2, 0 \leq n < w/2\} (\lambda = \text{LH, HL, HH}) \quad (5-16)$$

通过分析可知 3 个小波系数块的块内系数方差为

$$\sigma_{ij}^\lambda = \frac{1}{L/2 \times W/2} \sum_{m=0}^{L/2-1} \sum_{n=0}^{W/2-1} \left(|\mu^\lambda(mi, nj)| - \mu_{ij}^\lambda \right)^2 (\lambda = \text{LH, HL, HH}) \quad (5-17)$$

式中, μ_{ij}^λ 可定义为

$$\mu_{ij}^\lambda = \frac{1}{L/2 \times W/2} \sum_{m=0}^{L/2-1} \sum_{n=0}^{W/2-1} |\mu^\lambda(mi, nj)| (\lambda = \text{LH, HL, HH}) \quad (5-18)$$

则 3 个小波系数块的块内系数方差平均值 (即图像分割特征) 为

$$\sigma_{ij} = \frac{\sigma_{ij}^{\text{LH}} + \sigma_{ij}^{\text{HL}} + \sigma_{ij}^{\text{HH}}}{3} \quad (5-19)$$

(3) 利用 K 均值聚类算法对上述图像分割特征进行聚类 (分成两类), 聚类中心分别代表感兴趣区和背景, 即图像分割特征 σ_{ij} 大, 表示对应的空间域图像块属于感兴趣图像块。

(4) 将所有感兴趣图像块作为要处理的目标集合 S , 并将集合 S 按照 8 连通准则划分为满足如下条件的感兴趣图像块连通集 S_0, S_1, \dots, S_{P-1} , 其中 P 为连通集个数。

- ① $\bigcup_{k=0}^{P-1} S_k = S$ 即所有的感兴趣图像块连通集组成了整个目标集合;
- ② $S_i \cap S_j = \emptyset (i \neq j, 0 \leq i, j \leq P-1)$, 其中 \emptyset 为空集;
- ③ $S_k (0 \leq k \leq P-1)$ 是一个感兴趣图像块连通集;



④ S_i 与 S_j 非连通, 其中 $i \neq j, 0 \leq i, j \leq P-1$ 。

假设感兴趣图像块连通集 S_k 中共包含有 M_k 个感兴趣图像块, 且 $p_i^{(k)} (0 \leq i < M_k (L \times W))$ 为连通集 S_k 中第 i 个像素, 并记像素 $p_i^{(k)}$ 的坐标为 $(x_i^{(k)}, y_i^{(k)})$ 。取变量,

$$\begin{aligned} l_k &= \min \{x_0^{(k)}, x_1^{(k)}, \dots, x_{M_k(L \times W)-2}^{(k)}, x_{M_k(L \times W)-1}^{(k)}\} \\ r_k &= \max \{x_0^{(k)}, x_1^{(k)}, \dots, x_{M_k(L \times W)-2}^{(k)}, x_{M_k(L \times W)-1}^{(k)}\} \\ t_k &= \min \{y_0^{(k)}, y_1^{(k)}, \dots, y_{M_k(L \times W)-2}^{(k)}, y_{M_k(L \times W)-1}^{(k)}\} \\ b_k &= \max \{y_0^{(k)}, y_1^{(k)}, \dots, y_{M_k(L \times W)-2}^{(k)}, y_{M_k(L \times W)-1}^{(k)}\} \end{aligned}$$

则可以定义一个矩形 $R = \{(l_k, t_k), (r_k, b_k)\}$, 并称矩形 R 为感兴趣图像块连通集 S_k 的连通矩形, 记为 $R(S_k)$ 。再取变量 $A = |r_k - l_k| \times |b_k - t_k|$, 并称 A 为感兴趣图像块连通集 S_k 的连通面积, 并记为 $A(S_k)$ 。

对所有感兴趣图像块集 S 进行连通集划分, 将得到一个连通集 $\{S_k\} (0 \leq k \leq P-1)$, 而只要在感兴趣图像块连通集集合 $\{S_k\} (0 \leq k \leq P-1)$ 中选取连通面积较大的连通集 Ω_i , 则相应的连通矩形 $R(\Omega_i)$ 即为所要提取的感兴趣区。

文献[139]提取感兴趣区域的颜色、纹理和位置等特征来表示感兴趣区域的综合特征, 感兴趣区域的颜色特征用该区域的颜色均值和方差表示; 感兴趣区域的纹理特征可以采用小波域的子带能量特征进行描述; 位置特征采用感兴趣区域的重心的归一化坐标表示。

实验结果表明, 当图像中有明显的感兴趣区域时 (特别是背景简单的图像), 该方法能够更加准确地查找出用户所需内容的图像, 具有较好的检索精度。

5.4 基于颜色和熵的感兴趣区域图像检索

陆伟等^[140]提出利用图像的颜色和熵提取感兴趣区域, 并与情感特征建立有效的联系, 再通过映射从图像库中选取符合条件的图像, 最终提供检索结果。



1. 利用颜色特征提取感兴趣区域

图像颜色直接影响着人们对图像产生的印象并且易于提取,可以利用颜色的差异可以查找图像的感兴趣区域。颜色直方图是颜色整体的统计特性,不同的图像可能产生相同的直方图,相似的图像之间的直方图可能差异很大,因此陆伟等采用颜色平均值的差异性来提取感兴趣区域,并且采用 RGB 颜色模型空间。颜色由 R、G、B 分量表示,使用水平和垂直两种分块方法,在这两个方向上提取颜色特征,得到 R、G、B 值在这两种分块模型中分割区域的平均分布,如图 5-3 所示。

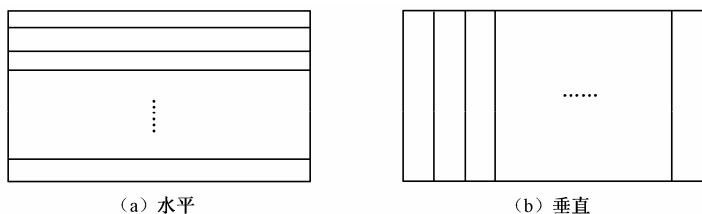


图 5-3 分块模型

每幅图像按水平和垂直被平均分为 L 个区域,每个区域的像素数 N 为 $H \times W / L$,其中 H 和 W 分别表示图像的高度和宽度。设 C 表示像素的 R、G、B 值, C_{RA} (Regional Average) 表示这 L 个区域的颜色平均值,则

$$C_{RA} = \sum_{i=1}^N C(i) \quad (5-20)$$

从这 L 个区域中按 R、G、B 颜色特征的差异选取 $L/2$ 个关键区域作为感兴趣区域,将这些感兴趣区域的颜色均值作为映射到情感空间的底层特征。颜色特征差异提取感兴趣区域的算法如下所述。

(1) 计算出 L 个区域中每个区域颜色的平均值 C_{RA} ,比较各区域之间 C_{RA} 的差别 D_{RA} ,即

$$D_{RA}(i) = \left| C_{RA}(i) - \sum_{j=1}^L C_{RA}(j) / L \right| (1 \leq i \leq L) \quad (5-21)$$



(2) 按照 D_{RA} 的大小选出 D_{RA} 最大的 $L/2$ 个区域作为感兴趣区域。

D_{RA} 的值越大, 表示此区域的颜色平均值与其他区域的颜色平均值差别越大, 即此区域与其他区域颜色的对比反差越大。利用这种方法选取的感兴趣区域具有较好的代表性, 符合颜色差异影响感性认识的条件。因此, 将感兴趣区域的 C_{RA} 值作为感性图像检索的一组底层特征, 能够降低映射算法的复杂度, 减少检索时底层图像特征的冗余度, 保证映射的有效性和准确性。这样从图像中提取出来的颜色特征总共应有 $3C \times L / (2C_{RA}) \times 2$ 个, 提取出的 $3L$ 个颜色特征组成一组向量, 令其为 $V_C = \{V_{C1}, V_{C2}, \dots, V_{C3L}\}$, 最终将这 $3L$ 个图像的颜色特征按照各自的权值 $W_C = \{W_{C1}, W_{C2}, \dots, W_{C3L}\}$ 映射到情感特征空间。

2. 利用信息熵提取感兴趣区域

图像的信息熵体现了图像灰度分布的统计特性, 利用信息熵来描述图像的特征, 降低了提取图像特征的复杂性, 使得图像特征的维数大大降低。信息熵定义为

$$E = - \sum_{m=0}^{255} p_m \log(p_m) \quad (5-22)$$

式中, P_m 为具有灰度 m 的像素在图像中出现的概率。

利用熵提取感兴趣区域的方法与利用颜色提取感兴趣区域的方法类似, 仍然按照水平和垂直两种分块模型对图像进行分块, 对划分出来的 L 个区域分别求得它们的熵。熵越大, 则此区域的信息量越丰富, 对人产生主观感觉的影响也就越大。因此, 按照分块后个区域熵值的大小, 从这 L 个区域中挑选出 $L/2$ 个熵最大 E_{RA} 的区域作为感兴趣区域, 再将感兴趣区域的熵作为映射的另一组底层特征, 得到 $3C \times L / (2E_{RA}) \times 2$ (分块模型) 个熵特征, 即 $3L$ 个熵特征 $V_e = \{V_{e1}, V_{e2}, \dots, V_{e3L}\}$, 同样按照各自的权值 $W_e = \{W_{e1}, W_{e2}, \dots, W_{e3L}\}$ 映射到情感特征空间。

颜色和熵能够充分反应图像的物理特征, 将图像感兴趣区域的颜色特征 V_c 和熵特征 V_e 按照权值 W_c 和 W_e 的结合形成检索系统的底层特征



$V=\{V_c, V_e\}$ 。每幅图像共有 $6L$ 个图像特征作为映射的底层特征, 前 $3L$ 个特征为颜色特征, 后 $3L$ 个特征为熵特征, 再通过神经网络的学习训练调整权值将图像特征映射到情感特征。

3. 基于感兴趣区域的感性图像检索

提取感兴趣区域的图像特征后, 要实现的是图像特征和情感特征之间相互的映射, 从图像特征到情感特征的转换需要一个映射的过程, 这种映射是一个非线性的过程。采用 BP 神经网络结构建立映射关系, 对系统进行训练学习调整权值实现图像特征到情感特征的映射。

检索的过程如下所述。

(1) 由用户输入代表情感特征的关键字, 对图像库中每幅图像进行感兴趣区域的提取。

(2) 通过神经网络的训练学习建立映射关系, 获得与此情感特征相匹配的图像。

(3) 由系统选择匹配程度最高的几幅图像显示给用户。由于感兴趣区域的提取减少了网络底层结点的维数, 检索便能够快速有效地进行。

实验结果分析表明, 单纯采用颜色和熵提取感兴趣区域进行检索时准确率相对较低, 同时利用颜色和熵提取感兴趣区域进行检索的平均准确率能够达到 77.163%。

5.5 本章小结

基于感兴趣区域的图像检索的研究能够有效地将图像和人类的视觉感知联系起来, 是基于内容的图像检索研究热点。本章讨论了 3 种典型的感兴趣区域的图像检索, 包括基于视觉显著度模型的图像检索、基于小波域感兴趣区域的图像检索, 以及基于颜色和熵的感兴趣区域的图像检索算法, 然而感兴趣区域的获取仍然受到图像分割算法的限制, 因此图像分割领域的研究必将推动图像检索的进一步发展。

第六章



基于兴趣点的图像检索

6.1 兴趣点概述

兴趣点是图像中灰度在 x 和 y 方向都有很大变化的一类特殊点，它们往往位于图像中视觉重要的部分，是认知心理学中的注视焦点，在其附近的局部区域内蕴含了丰富的图像显著细节信息。兴趣点作为一种重要的图像视觉特征点，它计算简单、信息含量高、匹配容易、数据量小等优点，已经在图像匹配、运动估计等许多领域得到了广泛的应用。

兴趣点通常被认为是角点（Corner Point），然而图像中的兴趣点除了是角点外，还可能是图像中纹理、形状和空间的显著点等^[141]。目前检测兴趣点的算法大致可分为 4 种，即基于图像边缘、基于模板匹配、直接从图像中提取兴趣点^[142]和小波兴趣点检测算法。

1. 基于边缘的兴趣点检测

该方法首先提取图像边缘，然后沿着边缘寻找曲率最大的点或拐点认为是兴趣点^[143]。文献[144]等在边缘上每一点的周围邻域内寻找是否有边的交点，若有交点则计算此点与交点的距离，根据距离判定是否是兴趣点。

基于边缘的兴趣点检测方法具有计算简便等优点，然而边缘检测算法往往受到噪声等影响，从而产生较大的误差，因此算法的鲁棒性不强。



2. 基于模板匹配的兴趣点检测算法^[145-146]

该方法需要先定义一些兴趣点模板,目标图像根据模板的角度、对称轴等进行特征点匹配,符合的则为兴趣点。

基于模板匹配的兴趣点检测算法检测比较准确,但往往需要相应的先验知识,因此算法的通用性略显不足。

3. 直接从图像中提取兴趣点的算法

文献[147]提出利用灰度差异来检测兴趣点,在固定窗口内计算某点与其近邻点的灰度自相关函数,若相关函数的最大值小于预定阈值,则认为是兴趣点。文献[148]提出根据图像灰度的自相关函数,计算 \hat{C} 矩阵的特征值,即

$$\hat{C} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \hat{I}_x^2 & \hat{I}_x I_y \\ \hat{I}_x I_y & \hat{I}_y^2 \end{bmatrix} \quad (6-1)$$

式中, \hat{I} 表示对 I 作高斯滤波。若 \hat{C} 矩阵的特征值较大时,则认为是兴趣点。文献[149]根据图像的梯度幅度与梯度方向变化率进行兴趣点检测,即

$$K = \frac{I_{xx}I_y^2 + I_{yy}I_x^2 - 2I_{xy}I_xI_y}{I_x^2 + I_y^2} \quad (6-2)$$

若 K 值较大可认为兴趣点。

4. 小波域兴趣点的算法

小波变换可以提取图像在不同尺度下的变化信息,从而得到小波域的图像兴趣点^[150]。对于一维信号函数 $f(x)$, 设 $\psi_{2^j}(x) = 2^j \psi(2^j x)$ 是 $\psi(x)$ 在尺度 2^j 的伸缩小波系, 设 N 是 $f(x)$ 的采样点数, 则 f 在 2^j 尺度的细节信号为 f 与平移伸缩小波的内积, 即

$$W_{2^j} f(n) = \langle f(u), \psi_{2^j}(u - 2^{-j}n) \rangle, 0 \leq n \leq 2^j N \quad (6-3)$$

可以在不同的尺度 2^j (即在尺度 $1/2, 1/4, \dots, 2^j, j \in Z$ 且 $j \leq -1$)



下分别利用式 (6-4) 计算其小波系数:

$$Wf = (W_{2^j} f)_{-J_{\max} \leq j \leq -1}, J_{\max} = \log_2 N \quad (6-4)$$

金字塔算法^[151]根据在尺度 2^{j+1} 的小波系数来计算尺度 2^j 的小波系数, 即

$$A_{2^j} f(n) = \sum_{k=-\infty}^{\infty} h(k-2n) A_{2^{j+1}} f(k), 0 \leq n < 2^j N \quad (6-5)$$

$$W_{2^j} f(n) = \sum_{k=-\infty}^{\infty} g(k-2n) A_{2^{j+1}} f(k), 0 \leq n < 2^j N \quad (6-6)$$

式中, h 是尺度离散滤波器, g 小波离散滤波器。

下面定义小波系数子集, 由式 (6-3) 可知, 计算在尺度 2^j 的小波系数 $W_{2^j} f(n)$ 所用的信号特征点集是 $\psi_{2^j}(u-2^{-j}n)$ 的空间支撑, 记为 $S(\psi_{2^j}(u-2^{-j}n))$ 。由金字塔方程可得 $W_{2^j} f(n)$ 和 $W_{2^{j+1}} f(n)$ 的空间支撑之间的关系, 即由式 (6-6) 得 $W_{2^{j+1}} f(n)$,

$$W_{2^{j+1}} f(n) = \sum_{k=-\infty}^{\infty} g(k-2n) A_{2^{j+2}} f(k), 0 \leq n < 2^{j+1} N \quad (6-7)$$

将式 (6-5) 代入式 (6-6) 中可得

$$W_{2^j} f(n) = \sum_{k=-\infty}^{\infty} (g(k-2n) \sum_{l=-\infty}^{\infty} h(l-2k) A_{2^{j+2}} f(l)), 0 \leq n < 2^j N \quad (6-8)$$

对于紧支撑小波, h 和 g 的空间支撑是相等的, 利用式 (6-7) 计算式 (6-8) 的支撑可得

$$S(\psi_{2^j}(u-2^{-j}n)) = \bigcup_{\{h/g(k-2n) \neq 0\}} S(\psi_{2^{j+1}}(u-2^{-j-1}k)) \quad (6-9)$$

滤波器 g 的大小为 $2p$ (对于 Haar 小波, $p=1$, Daubechies4 小波, $p=2$), $S(g(n))=[0, 2p-1]$, $\{k/g(k-2n) \neq 0\} = \{2n, \dots, 2n+2p-1\}$, 则给定小波系数的子集为

$$C(W_{2^j} f(n)) = \{W_{2^{j+1}} f(k), 2n \leq k \leq 2n+2p-1, 0 \leq n < 2^j N \quad (6-10)$$

由上述得到的小波系数子集只代表了 $2p$ 个信号特征点, 其中选择一个具有最大梯度值的信号特征点作为小波突出点。对每个小波特征子集系数按照式 (6-11) 进行显著小波特征值的计算,



$$\text{saliency} = \sum_{k=1}^{-j} |C^{(k)}(W_{2^j} f(n))|, 0 \leq n < 2^j N, -J_{\max} \leq j \leq -1 \quad (6-11)$$

就可以找到具有全局特征的小波兴趣点。

小波兴趣点提取算法可以扩充到二维数字图像 $I(M \times N)$ ，则小波系数子集为

$$C(W_{2^j}^d f(x, y)) = \{W_{2^{j+1}}^d f(k, l), 2x \leq k \leq 2x + 2p - 1, 2y \leq l \leq 2y + 2p - 1\}, \\ 0 \leq x < 2^j N, 0 \leq y < 2^j M, 1 \leq d \leq 3 \quad (6-12)$$

根据预设阈值对提取的小波系数子集进行阈值化，就可以提取小波变换的兴趣点。

下面将介绍两种典型的兴趣点图像检索算法，包括兴趣点环形空间特征检索算法和兴趣点扇形空间特征检索算法。这两种算法不仅提取了兴趣点的颜色特征，同时分析了兴趣点的空间分布性，具有较好的检索性能。

6.2 兴趣点环形空间特征

曾智勇^[141]提出了一种兴趣点环形空间特征的图像检索方法。该方法首先采用自适应非线性平滑滤波器进行图像增强，该滤波器基于各向异性热扩散方程，借用松弛迭代的思想，使得梯度幅度较大的边缘点（显著边缘）的幅度随迭代次数的增加而增强，而梯度幅度较小的边缘和噪声随着迭代而平滑。算法的具体实现过程如下所述。

(1) 计算梯度：

$$G_x(x, y) = 0.5[I(x+1, y) - I(x-1, y)] \quad (6-13)$$

$$G_y(x, y) = 0.5[I(x, y+1) - I(x, y-1)] \quad (6-14)$$

(2) 计算导热系数：

$$c(x, y) = \exp\left(-\frac{G_x^2(x, y) + G_y^2(x, y)}{2\alpha^2}\right) \quad (6-15)$$



(3) 由 $I^t(x, y)$ 迭代计算 $I^{t+1}(x, y)$, 其中, $I^0(x, y) = I(x, y)$ 。

$$I^{t+1}(x, y) = \frac{\sum_{i=-1}^1 \sum_{j=-1}^1 I^t(x+i, y+j) c^t(x+i, y+j)}{\sum_{i=-1}^1 \sum_{j=-1}^1 c^t(x+i, y+j)} \quad (6-16)$$

式中, $I(x, y)$ 为图像中像素点 (x, y) 处的幅度值; α 为决定梯度显著度的自由参数。经过自适应非线性滤波后, 运用式 (6-2) 提取显著兴趣点, 若 K 值大于阈值 λ , 则可认为是显著兴趣点。图 6-1 所示为算法迭代 5 次后提取的显著兴趣点。从图 6-1 中可以看出, 提取出的显著兴趣点确实是大都位于图像的物体上的显著边缘上。



图 6-1 文献[141]算法提取的显著兴趣点

由于兴趣点往往分布在图像的边缘上, 蕴含了丰富的颜色细节, 所以利用兴趣点局部区域的颜色特征来描述图像具有一定的合理性。颜色直方图是最简单而有效的颜色特征描述方法, 但颜色直方图丢失了颜色的空间分布状态信息, 因此, 对于颜色组合相似但分布截然不同的图像会造成误检。为此, 曾智勇等^[141]设计了一种基于显著兴趣点的环形颜色直方图, 将图像空间按照显著兴趣点的空间分布划分成一系列同心圆环, 然后以每个兴趣点为中心展开一个小窗口, 此窗口中的点应该落在



物体中, 这些小窗口内的点代表图像中的物体, 统计每个圆环内部显著兴趣点邻域的颜色直方图。由于这些兴趣点集中于物体的显著边缘上, 因此显著兴趣点邻域的颜色空间定位准确。

算法具体的操作步骤如下所述。

设 I 是任意一幅大小为 $M \times N$ 的图像, $(p_{i,j})_{M \times N}$ 是从图像上提取的显著兴趣点, $p_{i,j}$ 为显著兴趣点 (i,j) 处的灰度值, 设 $w = \{(x,y) | (x,y) \text{ 是显著兴趣点}\}$, $|w|$ 是显著兴趣点的个数, B_1, B_2, \dots, B_L 是图像量化后的颜色区域, $C = (\bar{x}, \bar{y})$ 为显著兴趣点的质心, 其中, \bar{x} 和 \bar{y} 分别为

$$\bar{x} = \frac{1}{|w|} \sum_{(x,y) \in w} x, \quad \bar{y} = \frac{1}{|w|} \sum_{(x,y) \in w} y \quad (6-17)$$

令 R 是以 C 为圆心的最大半径, 定义 $R = \max[(x - \bar{x})^2 + (y - \bar{y})^2]^{1/2}, x, y \in w$, 对于给定的一个正整数 N , 把 R 分为 N 等份, 然后以质心 C 为圆心, 以 $(j \times R) / N$ 为半径 ($1 \leq j \leq N$) 画圆可得到 N 个环, 则由内至外每个环和 w 的交点把 w 分为 w_1, w_2, \dots, w_N 。图 6-2 所示的是基于显著兴趣点的环形区域示意图。

设 $|w_j|$ 表示环形区间 j 内显著兴趣点的数目, 选取各环内以每个显著兴趣点为中心的 $\lambda \times \lambda$ 临域内的所有像数点的集合, 统计它们的颜色值获得第 j 个圆环的环形直方图 $H^j = \{h_i^j | 1 \leq i \leq L\}$, $\forall (x,y) \in w_j, (x',y') \in ((x,y), \lambda)$, $h_{p_{x',y'}}^j = h_{p_{x,y}}^j + 1$ 。

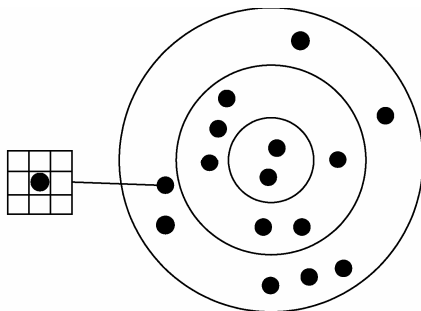


图 6-2 基于显著兴趣点的环形区域示意图



基于显著兴趣点的环形颜色直方图特征表征了兴趣点的颜色空间分布, 具有较好的检索性能, 然而算法的计算复杂度较高。为此, 文献[141]进一步提出了颜色空间信息熵特征。环形颜色直方图作为概率密度函数, 则基于显著兴趣点的颜色分布熵可表示为

$$E(H^j) = - \sum_{j=1}^N h_i^j \log_2(h_i^j) \quad (6-18)$$

显著兴趣点的颜色分布熵反映了具有相同颜色的小像素块在图像空间的分布特征, 由于人们往往对颜色块更感兴趣, 因而颜色分布熵更符合人的视觉特征。

基于兴趣点的环形空间特征检索算法把显著兴趣点作为图像中用户关注的视点线索, 通过它们找到视觉重要的局部区域, 利用局部区域的颜色特征及空间分布进行图像检索。该方法克服了颜色直方图没有位置信息的严重缺陷, 同时保证了检索算法对图像旋转、平移的鲁棒性。实验结果表明, 该方法实现简单, 与传统的基于兴趣点的图像检索方法相比具有更好的图像检索效率。

6.3 兴趣点扇形空间特征

孟繁杰等^[152]提出了一种将兴趣点局部分布特征和多示例学习相结合的图像检索方法, 图像按照兴趣点的分布划分成一系列扇形子区域, 提取出具有空间分布信息的兴趣点局部特征, 这些特征具有旋转、尺度和平移不变性, 具有较好的检索性能。

目前较多的兴趣点检测算法中, Harris 角点检测器具有较好的可重复性, 同时检测出的兴趣点对图像旋转、噪声和光照变化等都具有几何稳定性。然而, 该方法不具有尺度不变性。为此, 文献[152]采用尺度空间 Harris_Laplace 方法提取特征点。在提取兴趣点局部特征时, 为了包含兴趣点的空间分布信息, 可以对兴趣点进行扇形区域划分, 将所有兴趣点划分到同心等面积的扇形子区域中。通过这样的区域划分, 能够有



效地表示兴趣点从中心向四周各方向辐射分布的情况，并且计算简单，同时具有严格的几何不变性。图 6-3 所示的是扇形区域划分示意图。

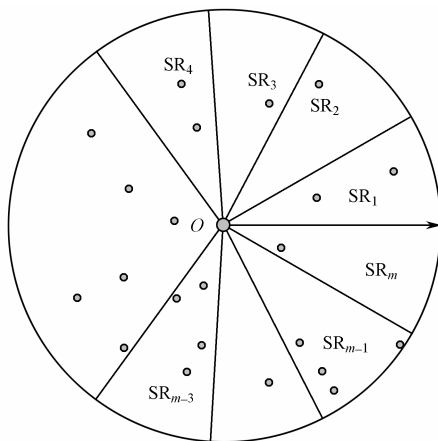


图 6-3 扇形区域划分

假设 (x, y) 是提取出一幅大小为 $W \times H$ 的图像上的兴趣点，令 P 为兴趣点集合， N 为兴趣点数目， $O(\bar{x}, \bar{y})$ 为兴趣点的质心，其中 \bar{x} 和 \bar{y} 分别为

$$\bar{x} = \frac{1}{N} \sum_{(x,y) \in P} x, \quad \bar{y} = \frac{1}{N} \sum_{(x,y) \in P} y \quad (6-19)$$

设 R 是以 O 为圆心的最大半径，即

$$R = \max[(x - \bar{x})^2 + (y - \bar{y})^2]^{1/2} \quad (6-20)$$

将以 O 为圆心、 R 为半径的圆形区域分割成 m 个面积相等的扇形区域。

若图像发生旋转后，图 6-3 的划分就有可能将相同的内容划分到不同序号的子区域中，因此在扇形区域划分前需要先对图像进行旋转归一化处理。对每个扇形子区域，选取其中所有兴趣点 δ 邻域内的像素，统计 HSV 空间颜色直方图 $h_k (1 \leq k \leq m)$ ，利用兴趣点的空间离散度表示每个子区域内兴趣点分布的疏密程度，第 k 个扇形子区域内兴趣点的空间离散度计算公式为

$$\text{dis}_k = \frac{1}{R} \left[\frac{1}{N_k} \sum_{(x,y) \in P_k} (x - \bar{x})^2 + (y - \bar{y})^2 \right]^{1/2} \quad (6-21)$$



式中, P_k 为第 k 个扇形区域内的兴趣点集合, N_k 为相应的兴趣点数目, (\bar{x}, \bar{y}) 为相应的兴趣点质心。空间离散度特征相似度利用高斯函数定义为

$$s(Q, I) = \sum_{k=1}^m \frac{N_k}{N} \exp[-(\text{dis}_k(Q) - \text{dis}_k(I))^2] \quad (6-22)$$

为了加强图像检索性能, 算法同时提取了兴趣点的颜色特征。多示例学习中的训练集由若干个具有概念标志的包组成, 每个包包含若干个没有概念标志的示例, 首先利用 EM—DD 算法^[153]实现对训练集的学习, EM—DD 算法是将期望最大 (EM) 算法与多样性密度 (Diverse Density, DD) 算法相结合提出的一种多示例学习算法, 然后进行多轮图像检索, 将图像按照相似度由大到小返回给用户。

实验分析表明, 将兴趣点和广义多示例学习相结合的检索性能优于传统兴趣点的典型检索算法。

6.4 本章小结

兴趣点是一种重要的图像视觉特征点, 具有计算简单、信息含量高和数据量小等优点。本章介绍了 4 种兴趣点检测算法, 同时重点介绍了兴趣点环形空间特征检索算法和兴趣点扇形空间特征, 这两种算法不仅分析了兴趣点周围区域的颜色特征, 同时分析了特征点空间分布性, 具有较好的检索性能。

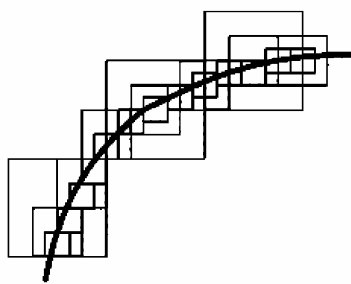
第七章



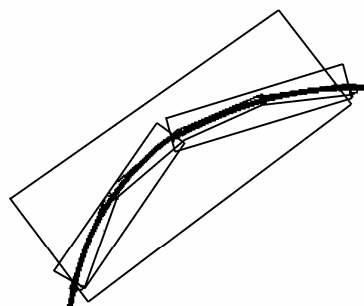
基于 Contourlet 变换的图像检索

7.1 Contourlet变换原理

小波变换具有较好的时频分析能力,能够反映图像奇异点的位置和特性,因此在图像处理领域获得较大的成功。小波变换能较好地刻画边缘造成的不连续性,却无法刻画出边缘本身的连续性,同时小波变换仅具有水平、垂直和对角 3 个方向的图像信息,这些特性直接影响了其在图像处理的应用。图 7-1 (a) 表示了用二维可分离小波来逼近图像中奇异曲线的过程。



(a) 小波变换的逼近曲线



(b) 理想变换的逼近曲线

图 7-1 小波变换和理想变换对曲线的描述

图 7-1 (b) 所示为理想变换的支撑区间表现为“长条形”,这种变换能充分利用原函数的集合正则性,达到用最少的系数来逼近奇异曲



线,基的“长条形”支撑区间实际上是“方向”性的一种体现,也称这种基具有“各向异性”,这种理想的变换就是“多尺度几何分析”。

在高维情况下,小波变换没有充分利用数据的几何特征,不能有效表征高维空间的奇异性,多尺度几何分析针对该问题进行了研究,Ridgelet 变换、Curvelet 变换、Contourlet 变换、Directionlet 变换和 Wedgelet 变换等^[154-159]相继被提出,已被成功应用于图像处理和模式识别等领域中。Contourlet 变换^[156]被认为是一种“真正”的图像二维表示方法,可以从不同方向表示图像纹理的多尺度几何信息,已在图像处理和纹理分析算法中体现了较强的优越性。

Contourlet 变换主要由两个步骤实现,即拉普拉斯金字塔分解 (Laplacian Pyramid, LP) 和方向滤波器组 (Directional Filter Bank, DFB) 滤波。首先,通过 LP 塔式分解把原始图像分解为低频子带和高频子带,高频子带再经过方向滤波器组分解为多个方向子带,对低频子带重复上述过程就可实现图像的 Contourlet 变换分解。图 7-2 所示为 Contourlet 变换的滤波器组结构。

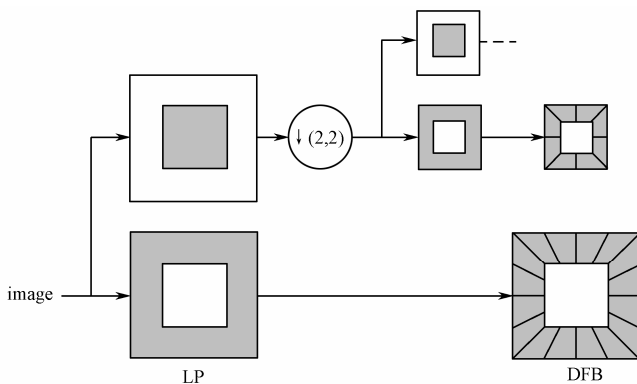


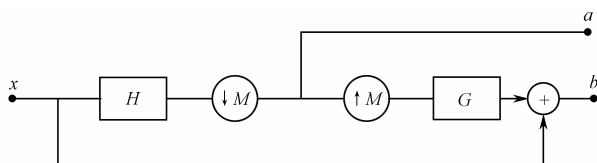
图 7-2 Contourlet 变换的滤波器组结构

7.1.1 拉普拉斯金字塔

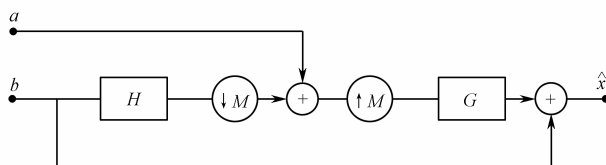
拉普拉斯金字塔 (LP)^[160]分解是一种多尺度分析方法,LP 分解可以把图像分解为低频子带和高频子带,分解和重构过程如图 7-3 所示。



图 7-3 (a) 中 H 和 G 分别被称为低通分解和合成滤波器, M 是抽样矩阵, x , a , b , \hat{x} 分别表示原始图像、低频子带、高频子带和重构图像, 对低频子带重复上述步骤就得到一个塔状结构, 因此称为 Laplacian 金字塔。



(a) Laplacian 金字塔分解



(b) Laplacian 金字塔重构

图 7-3 Laplacian 金字塔分解和重构

LP 金字塔滤波器分解可以捕捉奇异点, 产生原始信号的一个低通采样逼近及原始图像与低通预测图像之间的一个差值图像, 对得到的低通图像继续分解得到下一层的低通逼近和差值图像, 由此得到一系列的带通及下采样图像。由于 LP 每次尺度分解产生的带通信号只有一个, 因而有效避免了小波分析中所固有的频域混叠现象。

7.1.2 方向滤波器组

Bamberger 和 Smith^[16]于 1992 年提出了基于钻石形滤波器的方向滤波器组 (Directional Filter Bank, DFB), 可以实现对信号的多方向分解。然而, 钻石形 DFB 在多尺度迭代时的分解树规则比较复杂, 且需要对输入信号进行调制, 因此 Minh N. Do 提出了基于扇形的五株滤波器组 (Quincunx Filter Bank, QFB) 的方法, 信号通过对双通道 QFB 进行修剪操作, 可以实现方向滤波器的二叉树分解, 产生 2^l 个楔形的



频率划分。图 7-4 所示为 5 点梅花形采样矩阵组成的二维双通道滤波器组。

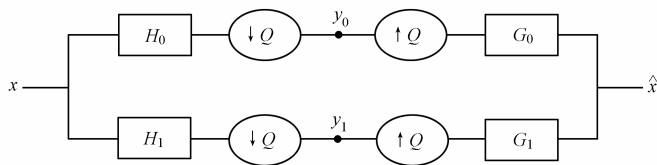


图 7-4 Quincunx 滤波器组

DFB 的分解结构可以使 Contourlet 变换可以在每个尺度上选择任意多的方向进行分解，图 7-5 所示为 DFB 在单尺度和多尺度下的频域分割图。其中，图 7-5 (b) 为在 3 个尺度下对图像频域的分割图，其中每个尺度的方向子带数目以 2 倍递增，即在 1, 2, 3 尺度下的方向子带数目分别为 2, 4, 8 个。

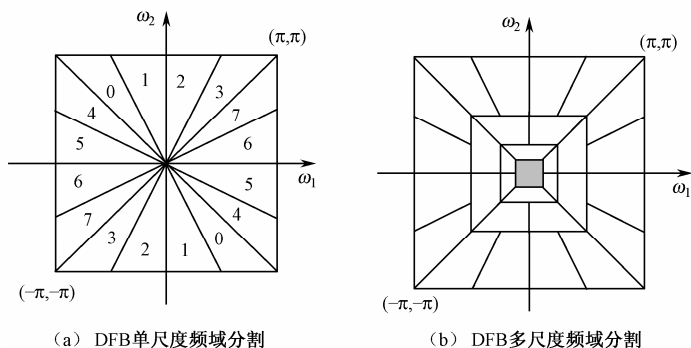


图 7-5 DFB 单尺度和多尺度分割

图 7-6 所示为 Zoneplate 图像进行 Contourlet 变换的 3 层分解示意图，其中 LP 和 DFB 分解都选择梯形结构的 pkva 滤波器，3 层分解的方向子带数分别为 0, 4 和 8。

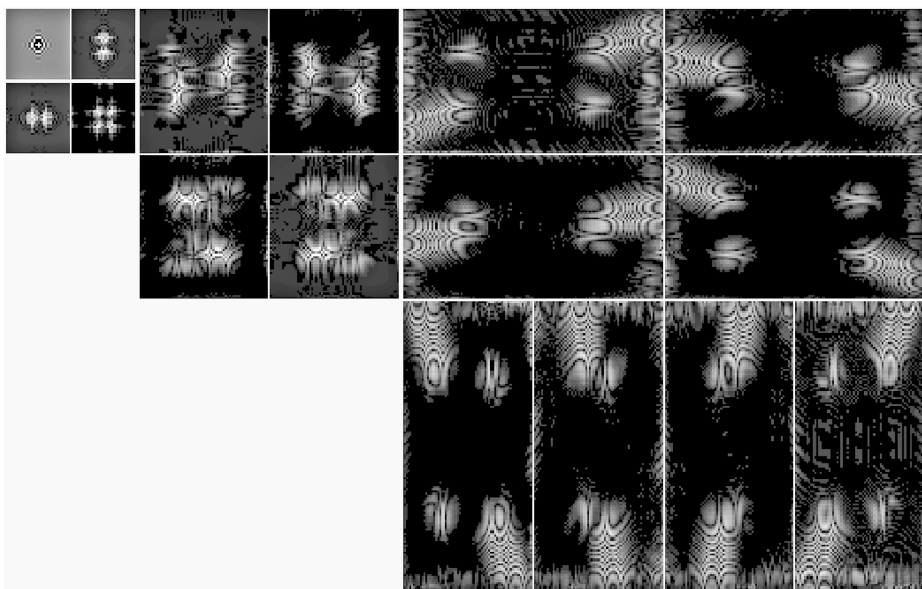


图 7-6 Zoneplate 图像的 Contourlet 变换示意图

7.2 基于Contourlet变换的检索算法

7.2.1 灰度—平滑共生矩阵

共生矩阵法在图像的纹理特征分析中具有广泛的应用,其中灰度—平滑共生矩阵可以有效地描述粗纹理,同时对图像旋转不敏感,具有极好的检索性能。纹理图像在 Contourlet 变换域的分析表明粗纹理空间能量集中在低频部分,而细纹理空间能量集中在高频部分,黄传波^[162]提出采用灰度—平滑共生矩阵法提取图像纹理的低频子带中的特征。设原始图像为 $f(x,y)$, $x=1,2,\dots,M$, $y=1,2,\dots,N$, 图像的灰度级为 $L=16$, 采用 3×3 矩阵对图像进行平滑滤波,得到平滑图像 $g(x,y)$, 则灰度—平滑共生矩阵元素 $A_{i,j}$ 的值为图像中 $f(x,y)=i$, $g(x,y)=j$ 的像素值。针对灰度—平滑共生矩阵进行归一化,得到其归一化矩阵 $\{p_{i,j}; i=1,2,\dots,L; j=1,2,\dots,L\}$, 可以计算灰度—平滑共生矩阵的反差、



熵、逆差距、能量、集群荫和集群突出共 6 个特征图像的纹理特征。

(1) 反差:

$$P_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} |i-j| p_{ij} \quad (7-1)$$

(2) 熵:

$$P_2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{ij} \log p_{ij} \quad (7-2)$$

(3) 逆差距:

$$P_3 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{p_{ij}}{1+|i-j|} \quad (7-3)$$

(4) 能量:

$$P_4 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{ij}^2 \quad (7-4)$$

(5) 集群荫:

$$p_5 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [(i-\mu_x) + (j-\mu_y)]^3 p_{ij} \quad (7-5)$$

(6) 集群突出:

$$p_6 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [(i-\mu_x) + (j-\mu_y)]^4 p_{ij} \quad (7-6)$$

式中, μ_x 为灰度平均值, $\mu_x = \sum_{i=0}^{L-1} i \left[\sum_{j=0}^{L-1} p_{ij} \right]$, μ_y 为平滑平均值,

$$\mu_y = \sum_{j=0}^{L-1} j \left[\sum_{i=0}^{L-1} p_{ij} \right]。$$

实验结果表明,在低频子带上抽取图像的灰度—平滑共生矩阵统计参数,可以充分利用低频子带上的丰富信息,能够取得较好的检索精度。由于低频子带上的图像经过多次抽样,低频子带与原图相比小得多,在低频子带上提取统计值会在一定程度上降低特征提取的时间复杂度。



7.2.2 广义高斯模型

传统 Contourlet 变换的方向子带系数具有非高斯性、高峰度和长拖尾等特点, 可以采用广义高斯分布进行描述, 广义高斯模型 GGD 的概率密度函数为^[28]

$$f(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta} \quad (7-7)$$

式中, α 为 GGD 模型的尺度参数, 它取决于随机变量的用来控制 $f(x; \alpha, \beta)$ 的宽度, β 为 GGD 模型的形状参数, 控制 $f(x; \alpha, \beta)$ 的形状, $\Gamma(x)$ 为伽玛函数: $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$ 。

GGD 模型中的参数 α 和 β 可采用矩估计和极大似然估计, 杨家红等^[163]提出了基于 Contourlet 广义高斯模型的纹理图像检索, 且采用矩估计方法实现。设 Contourlet 变换方向子带为 $M = \{x_1, x_2, x_3, \dots, x_n\}$, 其低阶矩可表示为

$$\begin{cases} m_1 = \int_{-\infty}^{\infty} |x| f(x; \alpha, \beta) dx = \frac{\alpha\Gamma(2/\beta)}{\Gamma(1/\beta)} \\ m_2 = \int_{-\infty}^{\infty} x^2 f(x; \alpha, \beta) dx = \frac{\alpha^2\Gamma(3/\beta)}{\Gamma(1/\beta)} \end{cases} \quad (7-8)$$

令 $g(\beta) = m_1^2 / m_2$, 即 $g(\beta) = \Gamma^2(2/\beta) / (\Gamma(1/\beta)\Gamma(3/\beta))$ 。为了计算 $g(\beta)$ 的值, 低阶矩可以由式 (7-9) 进行估计, 即

$$\begin{cases} m_1 = 1/n \sum_{i=1}^n x_i \\ m_2 = 1/n \sum_{i=1}^n x_i^2 \end{cases} \quad (7-9)$$

因此参数 α 和 β 可以由式 (7-10) 进行计算:

$$\begin{cases} \beta = g^{-1}(m_1^2 / m_2) \\ \alpha = m_1 \Gamma(1/\beta) / \Gamma(2/\beta) \end{cases} \quad (7-10)$$

为了分析 Contourlet 变换的子带系数分布是否符合广义高斯分布,



本文采用拟合度来描述广义高斯分布模型与实际参数的差异,拟合度可定义为

$$\text{fitness} = \frac{1}{N} \sum_{i=1}^N (h(x_i) - h^*(x_i))^2 \quad (7-11)$$

式中, $h(x_i)$ 表示实际系数直方图图形, $h^*(x_i)$ 表示通过参数估计确定的模型,而 N 则表示系数直方图的数量。实验采用 Brodatz 图像数据库中的 30 幅图像作为样本集,判断 Contourlet 变换子带系数是否符合广义高斯分布,表 7-1 为 Contourlet 变换拟合度参数 fitness 对照表。

表 7-1 Contourlet 变换拟合度参数 fitness 对照表

方向子带	均值	方差
0	0.0015	0.0003
1	0.0017	0.0006
2	0.0015	0.0004
3	0.0013	0.0004
4	0.0015	0.0005
5	0.0016	0.0005
6	0.0018	0.0007
7	0.0016	0.0004

由表 7-1 可知, Contourlet 变换的拟合度参数均值和方差非常小,说明子带系数均服从于广义高斯分布。文献[163]提出采用 KLD 距离作为两个子带之间相似性测度,设方向子带 G 和 H 的分布密度函数参数估计为 $p(G; \alpha_1, \beta_1)$ 和 $p(H; \alpha_2, \beta_2)$, 则子带 G 和 H 的 KLD 距离为

$$D(p(G; \alpha_1, \beta_1) / p(H; \alpha_2, \beta_2)) = \log \left(\frac{\beta_1 \alpha_2 \Gamma(1/\beta_2)}{\beta_2 \alpha_1 \Gamma(1/\beta_1)} \right) + \left(\frac{\alpha_1}{\alpha_2} \right)^{\beta_2} \frac{\Gamma((\beta_2 + 1)/\beta_1)}{\Gamma(1/\beta_1)} - \frac{1}{\beta_2} \quad (7-12)$$

检索实验表明,基于 Contourlet 变换 GGD 模型的检索性能优于传统的能量估计和熵估计方法。图 7-7 所示的是针对图 7-6 的第 3 层分解 8 个方向子带的 GGD 拟合,其中采用矩估计法进行参数估计。

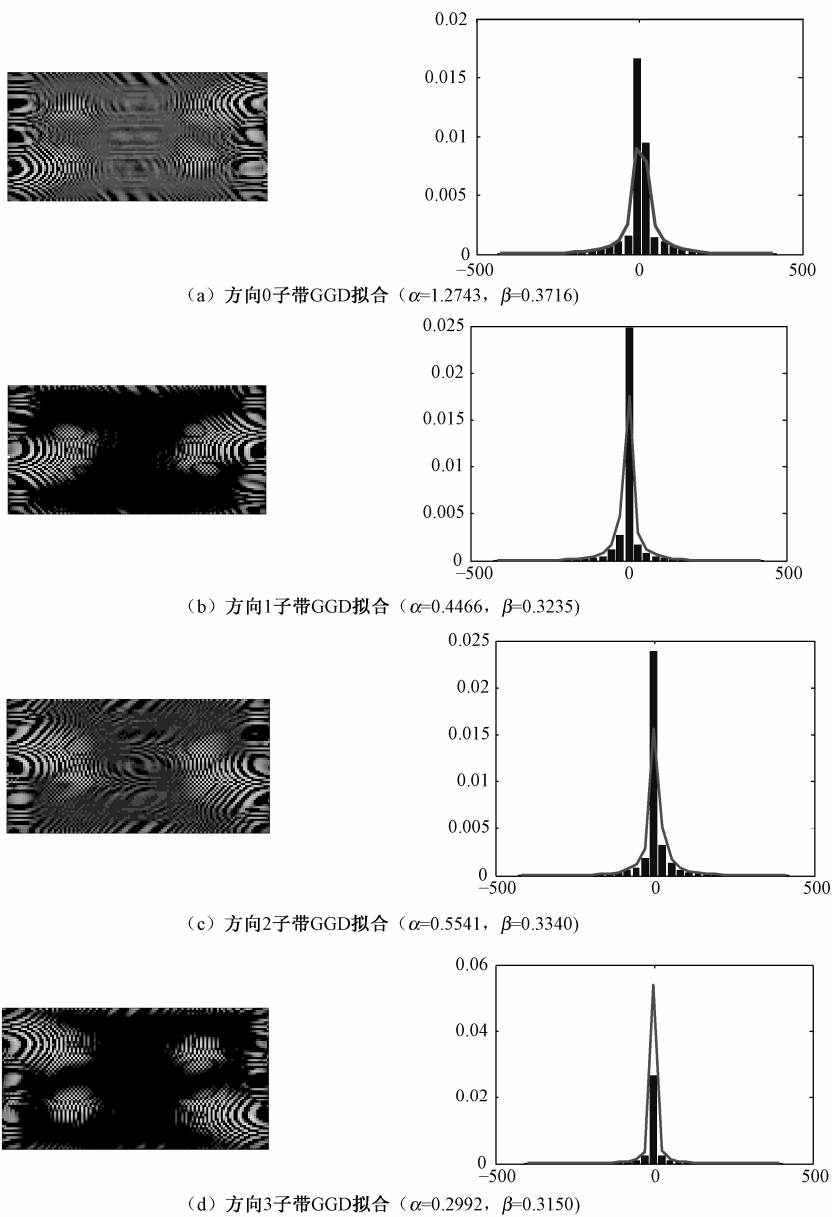
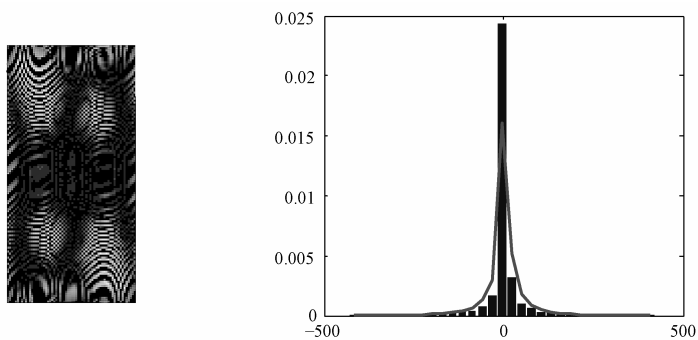
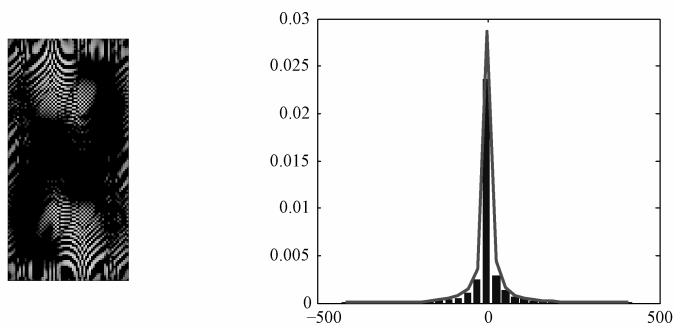


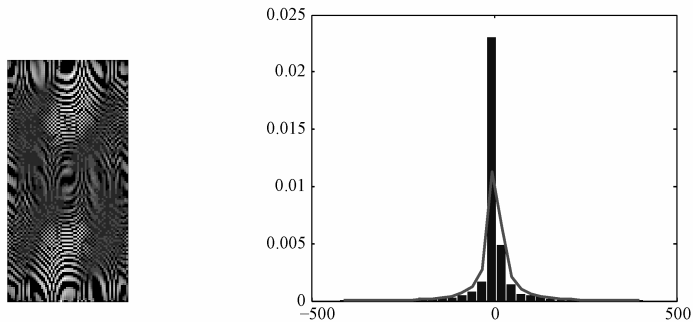
图 7-7 Contourlet 变换方向子带的 GGD 拟合



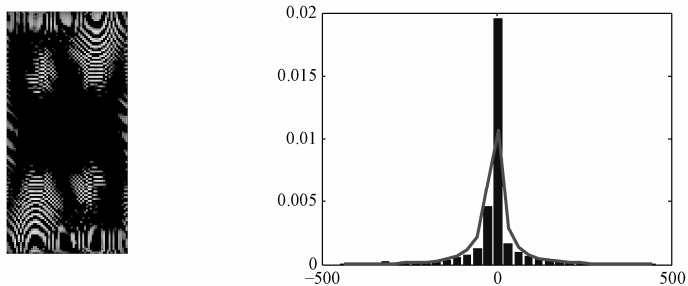
(e) 方向4子带GGD拟合 ($\alpha=0.5971$, $\beta=0.3338$)



(f) 方向5子带GGD拟合 ($\alpha=0.8137$, $\beta=0.3499$)



(g) 方向6子带GGD拟合 ($\alpha=0.6243$, $\beta=0.3365$)



(h) 方向7子带GGD拟合 ($\alpha=1.9771$, $\beta=0.3857$)

图 7-7 Contourlet 变换方向子带的 GGD 拟合 (续)



7.2.3 统计矩特征

Rao 等^[164]提出了基于 Contourlet 变换的统计矩方法进行检索, 首先将图像进行 Contourlet 变换分解, 然后计算 Contourlet 变换各方向子带的标准差, 即

$$\sigma_k = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (W_k(i, j) - \mu_k)^2} \quad (7-13)$$

式中, W_k 是 Contourlet 变换的第 k 个子带系数, μ_k 是 Contourlet 变换第 k 个子带系数的均值, 图像大小为 $M \times N$ 。图像特征可表示为 $f = [\sigma_1, \sigma_2, \dots, \sigma_L]$, 进一步可以将图像特征进行归一化:

$$\hat{f} = \frac{f - \mu_f}{\sigma_f} \quad (7-14)$$

式中, μ_f 和 σ_f 分别为特征向量的均值和方差。最后, 利用 Manhattan 距离作为特征向量间的距离测度。

实验结果表明, 基于 Contourlet 变换的统计矩方法在一定程度上优于 Gabor-Zernike 特征方法。

7.3 基于 NSCT 变换的检索算法

由于拉普拉斯金字塔分解和 DFB 都存在下采样和上采样操作, 因此 Contourlet 变换不具有平移不变性, 而平移不变性往往会影响其在图像处理中的应用。为了使 Contourlet 变换具有平移不变性, Zhou 等^[165]提出了一种非下采样 Contourlet 变换 (Nonsubsampled Contourlet Transform, NSCT)。NSCT 主要由两部分组成: 非下采样拉普拉斯金字塔分解 (Nonsubsampled Pyramid, NSP) 和非下采样方向滤波器组滤波 (Nonsubsampled Directional Filter Bank, NSDFB)。

NSP 采用一组双通道无下采样滤波器组 (NSFB), 其理想频率响应如图 7-8 (a) 所示, 该滤波器组没有下采样和上采样操作, 因此 NSP



具有平移不变性^[165-167]。图像经过 N 级 NSP 分解后, 可得到 $N+1$ 个与源图像具有相同尺寸大小的子带图像。NSP 理想重构的条件是满足 Bozout 恒等式, 即

$$H_0(z)G_0(Z) + H_1(z)G_1(Z) = 1 \quad (7-15)$$

式中, $H_0(z)$ 和 $G_0(z)$ 分别表示低通分解和合成滤波器, $H_1(z)$ 和 $G_1(z)$ 分别表示高通分解和合成滤波器。

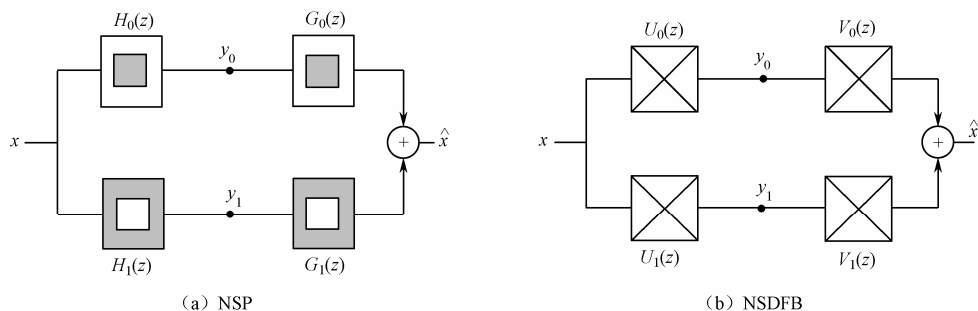


图 7-8 NSCT 的理想频率响应

NSDFB 采用的也是一组二通道无下采样滤波器组, 其理想频率响应如图 7-8 (b)所示, 其中 $U_0(z)$ 和 $U_1(z)$ 代表分解滤波器, $V_0(z)$ 和 $V_1(z)$ 代表重构滤波器。NSDFB 对每个尺度下的图像进行方向分解, 并且每个尺度下的方向子图大小都和原图大小相同。

非下采样金字塔结构保证了 NSCT 多尺度性, 而 NSDFB 保证了 NSCT 的多方向性, NSCT 变换的平移不变性可有效克服伪 Gibbs 现象, 因此 NSCT 在图像处理方面已获得了有效的应用。图 7-9 所示为 Zoneplate 图像的 NSCT 变换单层分解高频子带示意图 (参数为 $nlevels = [3]$)。

梁昕等^[166]提出了基于 NSCT 变换的 SAR 图像检索系统, 以系数幅度序列的均值 μ_i 、标准方差 σ_i 和三阶中心矩立方根 λ_i 作为 SAR 图像的纹理特征, 设某一尺度下的方向子带大小为 $M \times N$, 则

$$\mu_i = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N |C_i(x, y)| \quad (7-16)$$



$$\sigma_i = \sqrt{\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N (C_i(x, y) - \mu_i)^2} \quad (7-17)$$

$$\lambda_i = \sqrt[3]{\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N (C_i(x, y) - \mu_i)^3} \quad (7-18)$$

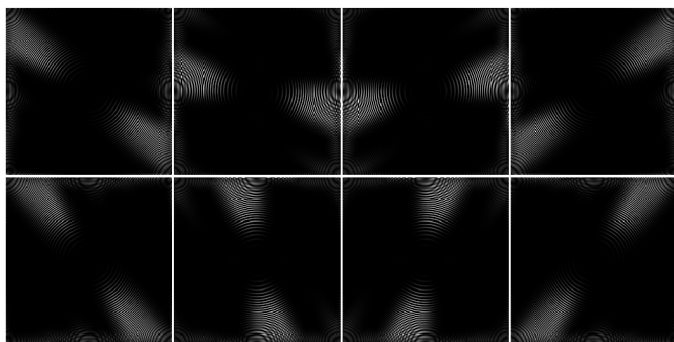


图 7-9 Zoneplate 图像的 NSCT 变换示意图

图像的特征向量可表示为 $F = [\mu_1, \sigma_1, \lambda_1, \dots, \mu_i, \sigma_i, \lambda_i, \dots, \mu_L, \sigma_L, \lambda_L]$, 其中 L 为子带数量。文献[166]在相似性度量时, 对均值、标准方差和三阶中心矩等进行了 3 种组合, 利用组合特征进行图像检索。

(1) 组合一: 均值、标准方差。

$$d_k = \left| \frac{\mu_k(x) - \mu_k(y)}{\sigma(\mu_k)} \right| + \left| \frac{\sigma_k(x) - \sigma_k(y)}{\sigma(\sigma_k)} \right| \quad (7-19)$$

(2) 组合二: 方差、三阶中心矩。

$$d_k = \left| \frac{\lambda_k(x) - \lambda_k(y)}{\sigma(\lambda_k)} \right| + \left| \frac{\sigma_k(x) - \sigma_k(y)}{\sigma(\sigma_k)} \right| \quad (7-20)$$

(3) 组合三: 均值、方差、三阶中心矩。

$$d_k = \left| \frac{\lambda_k(x) - \lambda_k(y)}{\sigma(\lambda_k)} \right| + \left| \frac{\sigma_k(x) - \sigma_k(y)}{\sigma(\sigma_k)} \right| + \left| \frac{\mu_k(x) - \mu_k(y)}{\sigma(\mu_k)} \right| \quad (7-21)$$

式中, $\mu_k(x)$, $\sigma_k(x)$, $\lambda_k(x)$ 和 $\mu_k(y)$, $\sigma_k(y)$, $\lambda_k(y)$ 分别为图像 x 和 y 的第 k 个子带的均值分量、标准分量和三阶中心矩分量, $\sigma(\mu_k)$, $\sigma(\sigma_k)$, $\sigma(\lambda_k)$ 为图像库中图像 x , y 各特征向量的标准差。

基于 SAR 图像的纹理特征检索实验表明, 采用基于 NSCT 算法的



方差和三阶中心矩的特征向量组合检索率优于其他两种方法，同时优于传统小波变换的方法。

7.4 圆对称Contourlet变换检索

为了克服传统 Contourlet 变换存在的频谱混淆现象，练秋生^[168-169]提出了构造具有平移不变性的圆对称轮廓波变换（Translation-Invariant Circular Symmetric Contourlet Transform, TICSCT）。TICSCT 与传统 Contourlet 变换相比，圆对称轮廓波变换的基函数的光滑度（即正则性）高，并且没有明显的频谱混淆现象，方向选择性更强。

TICSCT 由圆对称塔式分解（Circular Symmetric Pyramid Decomposition, CSPD）和非抽样方向滤波器组（Undecimated Directional Filter Banks, UDFB）两部分实现。图 7-10 所示的是 TICSCT 的原理图，其中 CSPD 将图像分解为多个不同分辨率的细节子带和一个低频逼近子带，而 UDFB 再将每个细节子带分解为多个方向子带。图 7-11 中 CSPD 的两个圆对称滤波器 L_0 和 H_i 满足以下重构条件：

$$L_0(\omega) = 0, \text{ 若 } |\omega| = \sqrt{\omega_1^2 + \omega_2^2} \geq \pi/2 \quad (7-22)$$

$$H_i^2(\omega) + \frac{1}{4} L_0^2(\omega) = 1 \quad (7-23)$$

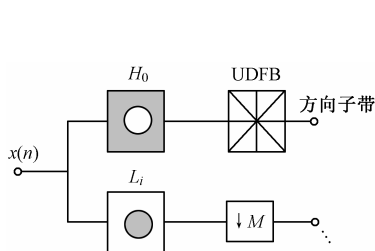


图 7-10 TICSCT 原理图

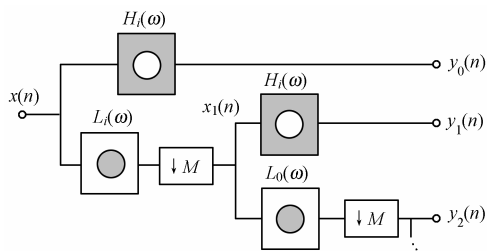


图 7-11 CSPD 原理图

尽管 CSPD 中存在下抽样操作，但只要低通滤波器的频域支撑区是半径为 $\pi/2$ 的圆形区域，CSPD 则具有平移不变性，同时 UDFB 无下



采样过程, UDFB 也具有平移不变性, 因此由具有平移不变性 CSPD 和 UDFB 实现的 TICSCT 也同样具有平移不变性。

文献[168]分析了 TICSCT 系数边缘分布具有非高斯性、高峰度和长拖尾等特点, 可以采用广义高斯模型描述, 利益矩估计方法描述各方向子带的特征, 同时融合局部二进制模式 (LBP) 特征综合描述图像的纹理特征, 由于两种特征是分别在空域和变换域提取的, 它们具有一定的互补性。而视觉感受明显不同的纹理图像可能具有相同的子带统计特性, 因此融合两种纹理特征的图检索具有较好的检索效果。

7.5 复数 Contourlet 变换检索算法

复数轮廓波变换具有良好的平移不变性, 已被广泛应用于图像处理领域中, 陈新武等^[170]提出了基于复数轮廓波变换纹理图像检索系统。为了有效降低变换的冗余度, 保持较高的时移不变性并克服变换域系数的振荡特性, Chen 等^[171]提出了复数轮廓波变换, 其构造方法是采用双树复小波变换级联临界下采样方向滤波器, 这样可以保持较低的冗余度并具有良好的时频局部化特征。复数轮廓波变换已被成功应用于图像融合和图像消噪^[172]。

复数轮廓波变换包括两个组成部分: 首先采用 Kingsbury^[124]提出的 DTCWT 实现多尺度分解, 在尺度 2^j 上的细节子空间 W_j 包含一个双树结构, 得到 6 个方向子带, 在位置 $x = (x_1, x_2)$ 上产生的系数可以用式 (7-24) 进行描述:

$$\{y_{m,n,p}^i(x)\}_{m,x \in Z^2} \quad (7-24)$$

式中, $p \in \{1, 2\}$, 表示小波系数的实数部分和虚数部分; $i \in \{1, 2, 3\}$, 与 p 结合在一起, 可以用 12 个子带来描述 6 个不同的方向信息; $m = (m_1, m_2)$, $n = (n_1, n_2)$ 表示位置偏移。复数小波变换的输出可以构成一个准 Hilbert 变换对, 因而细节子空间 W_j 具有近似的移不变性质。为了进一步提高复小波变换的方向分辨力, 可以将 DFB 级联在其细节子带



上, 将细节子空间进一步分解, 每一个子带可以被扩展至 2^{l_j} 个子带空间, 即

$$\mu_{k,m,p,d}^{i,l_j}(x) = \sum_{m \in Z^2} g_k^{l_j} [m - S_k^{l_j} n] \gamma_{m,p,d}^i(x), i \in \{1, 2, 3\} \quad (7-25)$$

式中, $\mu_{k,m,p,d}^{i,l_j}(\cdot)$ 表示尺度 2^j 上的方向子空间族 $W_{j,k}^{l_j}, k \in (1, 2, \dots, 2^{l_j})$, $g_k^{l_j}(\cdot)$ 表示重构滤波器的冲激响应。在这种情况下, 每一个方向子空间 $W_{j,k}^{l_j}$ 包含一个复双树框架, 这种基于双树复小波和方向滤波器架构的复轮廓波变换称为“复数轮廓波变换 (Complex Contourlet Transform, CCT)”。图 7-12 所示的是复数轮廓波变换的结构图。

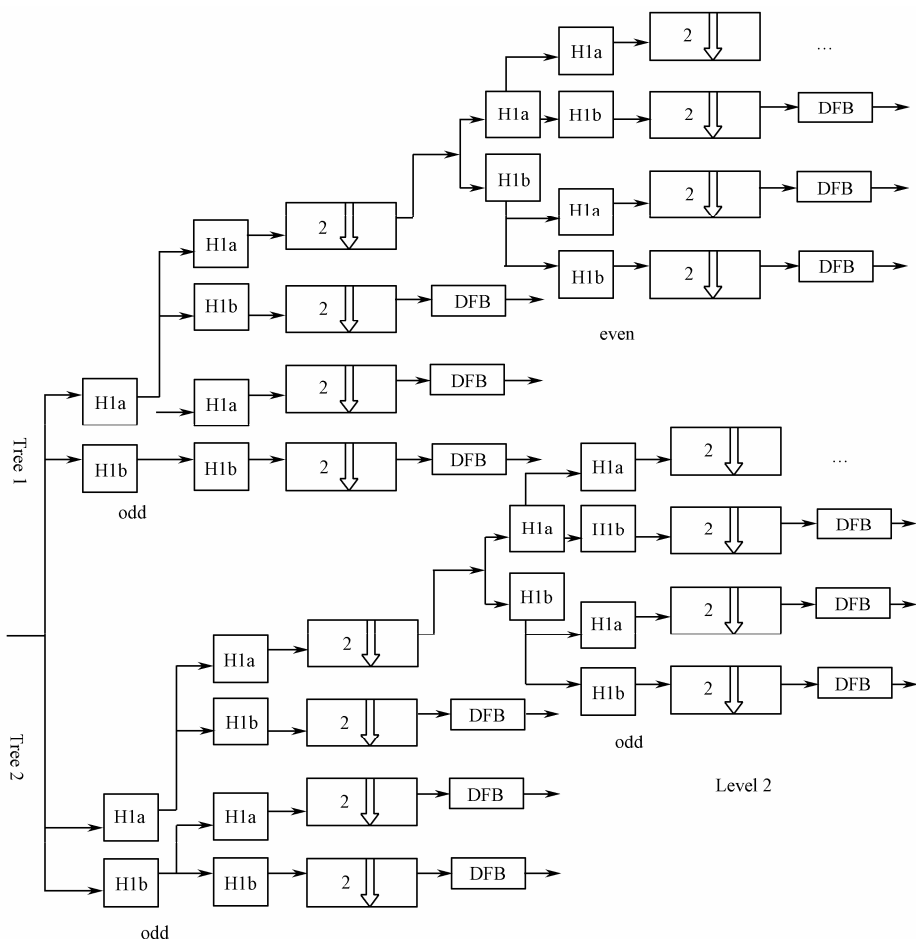


图 7-12 复数轮廓波变换的结构图



陈新武等^[164]采用复轮廓波变换各方向子带的能量和标准差作为图像纹理的特征向量,同时利用 Canberra 距离作为相似度衡量标准。

实验选取 Brodatz album 纹理图像数据库中的 109 幅纹理图像,将复数轮廓波变换与 Contourlet、NSCT 进行了比较。实验结果表明,基于复数轮廓波变换的检索系统比其他变换的检索系统都具有较高的检索率,同时随着分解尺度数的增加,检索率呈上升趋势。

7.6 本章小结

Contourlet 变换具有良好的时频分析能力和的方向表达能力,已在图像检索算法中体现了较强的优越性。本章介绍了传统 Contourlet 变换的原理,然后重点介绍了 3 种基于 Contourlet 变换的图像检索算法,最后介绍了近年来的最新研究成果,其中包括基于 NSCT 变换的检索算法、圆对称 Contourlet 变换检索算法和复数轮廓波变换检索算法,这些算法的提出在一定程度上促进了多尺度几何分析在图像检索领域的应用和发展。

参 考 文 献



- [1] Ravela S, manmatha R. Image retrieval by Appearance[C]. In Proceedings of SIGIR'1997, Philadephia PA, USA ,1997: 278-285.
- [2] 孙君顶, 赵珊. 图像低层特征提取与检索技术[M]. 北京: 电子工业出版社, 2009.
- [3] Flickner M., Sawhney H., Niblack W., et al. Query by image and video content: the QBIC System[J]. IEEE Computer, 1995, 28(9):23-32.
- [4] Pentland A., Picard R., Sclaroff S. Photobook: tools for content-based manipulation of image database[C]. Proc. SPIE, Feb. 1994, 2185: 34-47.
- [5] Picard R. W., Kabir T. Finding similar Patterns in large image databases[C]. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1993, 5: 162-164.
- [6] Smith J. R., Chang S. F. VisualSEEK: a fully automated content-based image query system[C]. Proc. ACM Multimedia, Nov. 1996, 87-98.
- [7] 王艳娟. 基于感兴趣区域的图像检索技术研究[D]. 阜新: 辽宁工程技术大学, 2007.
- [8] 庄越挺. 智能多媒体信息分析与检索的研究[D]. 杭州: 浙江大学, 1998.
- [9] 李国辉, 曹莉华, 柳伟等. 多媒体信息查询和检索系统 MIRC[J]. 小型微型计算机系统, 1999, 20(9): 672-676.
- [10] 焦李成, 张向荣, 侯彪等. 智能 SAR 图像处理与解译[M]. 北京:



- 科学出版社, 2008.
- [11] 崔江涛. 高维索引技术中向量近似方法研究[D]. 西安: 西安电子科技大学, 2005.
- [12] Ishikawa Y., Subramanya R., Faloutsos C. MindReader: Querying databases through multiple examples[C]. Proc. VLDB, Aug, 1998:98-109.
- [13] Cox I. J., Miller M. L., MinLka T. P., et al. The Bayesian image retrieval system, PicHunter: theory, implementation and Psychophysical experiments[C]. IEEE Trans. On Image Processing, 2000, 9(1):20-37.
- [14] Nastar C, Mitschke M, Meihac C. Efficient Query Refinement for Image Retrieval[C]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Santa Barbara, California, IEEE Computer Society, 1998: 547-552.
- [15] Zhang C., Chen T. An active learning framework for content-based information retrieval.[C] IEEE Trans. On Multimedia, 2002, 4(2):260-268.
- [16] Qian F., Li M. J., Zhang L., et al. Gaussian mixture model for relevance feedback in image retrieval[C]. IEEE conf. on ICME, 2002:229-232.
- [17] 章毓晋. 基于内容的视觉信息检索[M]. 北京: 科学出版社, 2003.
- [18] Smith J. R., Chang S. F. Automated binary texture feature sets for image retrieval[C]. In Proc. ICASSP, Atlanta, 1996, 4:2239-2242.
- [19] 沈兰荪. 图像检索与压缩域处理技术的研究[M]. 北京: 人民邮电出版社, 2008.
- [20] 李勇. 基于内容的图像检索技术研究[D]. 长春: 吉林大学. 2009.
- [21] Ma W. Y., Manjunath B. S. Edge flow: a framework of boundary detection and image segmentation[C]. IEEE Int. Conf. on Computer



- Vision and Pattern Recognition, 1997, 744-749.
- [22] Swain M. J., Ballard D. H. Color indexing[J]. International Journal of Vision Computer Vision, 1991, 7(1): 11-32.
- [23] 庄越挺. 网上多媒体信息分析与检索[M]. 北京: 清华大学出版社, 2002.
- [24] 盛骤, 谢式千, 潘承毅. 概率论与数理统计[M]. 北京: 高等教育出版社, 1989.
- [25] Zhang D. S., Lu G. J. Evaluation of similarity measurement for image retrieval[C]. IEEE Int. Conf. Neural Networks & Signal Processing, Nan Jing, China, 2003, 928-931.
- [26] Santini S., Jain R. Similarity match[J]. IEEE Trans. On PAMI, 1996, 18(9): 946-958.
- [27] Kullback S. Information theory and statistics[M]. Dover, New York, NY, 1968.
- [28] Minh N. D., Martin Vetterli. Wavelet-Based texture retrieval using generalized gaussian density and kullback-leibler distance[J]. IEEE Transactions on Image Processing, 2002, 11(2): 146-158.
- [29] Puzicha J., Hufnann T., Buhmann J. Non-parametric similarity measure for unsupervised texture segmentation and image retrieval[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 1997, 267-272.
- [30] Zhang D. S. Image Retrieval Based on Shape[D]. PhD Thesis, Monash University, March, 2002.
- [31] Arnold W. M., Marcel W., Simone S., et al. Content-based image retrieval at the end of the early years[J]. IEEE Trans. On PAMI, 2000, 22(12): 1349-1379.
- [32] Mehtre B. M., Kankallalli M. S. Narasimhalu A. D. Color matching



- for image retrieval[J]. Pattern Recognition Letters, 1995, 16:325-331.
- [33] Cha G. H., Chung C. W. The GC-tree: a high-dimensional index structure for similarity search in image database[J]. IEEE Trans. on Multimedia, 2002, 4(2): 235-247.
- [34] Berm P., Shapiro G. A flexible image database system for content-based retrieval[J]. Computer Vision and Image Understanding, 1999, 75(1/2):175-195.
- [35] Bechmann N., Kriegel H. P. The R*-tree: an efficient and robust access method for points and rectangles[C]. Proc. of the ACM SIGMOD conference on the Management of Data, 1990, 322-331.
- [36] White D. A., Jain R. Similarity indexing with the SS-tree[C]. Proc. of International conf. on Data Engineering, 1996, 516-523.
- [37] Hsieh I. S., Fan K C. Multiple classifiers for color flag and trademark image retrieval[J]. IEEE Trans. On Image Processing, 2001, 10(6):938-950.
- [38] Faloutsos C., Lin K. Fastmap: A fast algorithm for indexing data-mining and visualization of traditional and multimedia datasets[C]. In Proc. Of ACM SIGMOD, 1995, 163-179.
- [39] Chandrasekaran S., Manjunath B. S., Wang Y. F., et al. An eigenspace update algorithm for image analysis[J]. Graphical Models and Image Processing, 1997: 689-696.
- [40] Salton G., McGill M. J. Introduction to modern information retrieval[M]. New York: McGraw-Hill Book Company, 1982.
- [41] Guttman A. R-trees: A dynamic index structure for spatial searching[C]. In Proc. ACM SIGMOD, 1984, 6:47-57.
- [42] Charikar M., Chandra C., Feder T., et al. Incremental clustering and dynamic information retrieval[C]. In Proc. Of the 29th Annual ACM



- Symposium on Theory of Computing, 1997, 626-635.
- [43] Zhang H. J., Zhong D. A scheme for visual feature based image retrieval[C]. In Proc. SPIE Storage and retrieval for image and video database, 1995.
- [44] Jain A. K., Vailaya A. Shape-based retrieval: a case study with trademark image databases[J]. Pattern Recognition, 1998, 31(9): 1369-1390.
- [45] Castleman K. R. 著, 朱志刚等译. 数字图像处理[M]. 北京: 电子工业出版社, 1998.
- [46] 章毓晋. 图像工程上册—图像处理与分析[M]. 北京: 清华大学出版社, 1999.
- [47] 周明全, 耿国华, 韦娜. 基于内容图像检索技术[M]. 北京: 清华大学出版社, 2007.
- [48] Stricker M, Orengo M. Similarity of color images[C]. In: Proceedings of SPIE Storage and Retrieval for Image and Video Databases III, San Jose, CA: Feb ,1995, 2420:381-392.
- [49] 孙君顶, 崔江涛, 毋小省等. 基于颜色和形状特征的彩色图像检索方法[J]. 中国图像图形学报, 2004,9(7):820-827.
- [50] 孙君顶, 武学东, 周利华. 基于颜色和形状的图像检索[J]. 计算机科学, 2004, 31(5):180-183.
- [51] 张海成, 彭晓明, 刘娜等. 基于主色调的广义直方图图像检索方法[J]. 空军雷达学院学报, 2005, 19(3):55-58.
- [52] 洪清启, 王备战, 董槐林等. 一种改进的基于广义直方图的图像检索方法[J]. 四川大学学报(工程科学版), 2007, 39(5):292-296.
- [53] J.Han, K.K, Ma. Fuzzy color histogram and its use in color image retrieval[J]. IEEE Transactions on Image Processing, 2002, 11(8) : 944-952.



-
- [54] Ju H, Kai-Kuang Ma. Fuzzy color histogram: an efficient color feature for image indexing and retrieval[C]. Proceedings of Acoustics, Speech and Signal Processing, ICASSP '00. 2000, 4:201-204.
- [55] 王伟, 徐玮, 熊志辉. 基于模糊直方图的两阶段相似视频的自动鉴别和检索[J]. 小型微型计算机系统, 2007, 8(8):1477-1481.
- [56] 黄朝兵, 余胜生, 周敬利. 基于多邻域统计矩直方图的彩色图像检索[J]. 小型微型计算机系统, 2005, 26(6):1061-1064.
- [57] Zachary J. M. An Information Theoretic Approach to Content Based Image Retrieval[D]. Louisiana State University and Agricultural and Mechanical College, Phd. Thesis, 2000, 45-62.
- [58] 孙君顶. 基于内容的图像检索技术研究[D]. 西安: 西安电子科技大学, 2005.
- [59] 沈云涛, 郭雷. 基于颜色视觉关注度的图像检索[J]. 计算机工程与应用, 2005, 41(31), 58-60.
- [60] 丁明军, 姚明海, 朱丽霞. 基于广义平均直方图的图像检索方法[J]. 科技通报, 2009, 25(1): 93-97.
- [61] 刘松堂, 吴钢. 基于广义直方图均衡的图像增强新方法[J]. 电光与控制, 2010, 17(3):12-15.
- [62] Smith J. R., Chang S. F. Tools and techniques for color image retrieval[C]. In Proc. of SPIE: Storage and Retrieval for Image and Video Database. 1995, 2670: 426-437.
- [63] Pass G., Zabini R., Miller J. Comparing Images Using Color Coherence Vectors[C]. In ACM International Conference on Multimedia. Boston, MA, 1996, 65-73.
- [64] Huang J. Color-Spatial Image Indexing and Applications[D]. Phd thesis, Cornell University, New York, USA, 1998.
- [65] 孙君顶, 张喜民, 崔江涛等. 一种新的基于颜色和空间特征的图像



- 检索方法[J]. 计算机科学, 2005, 32(6):158-160.
- [66] Cinque L., Levialdi S., Olsen K. Color-based image retrieval using spatial-chromatic histograms[C]. In Multimedia Computing and System, IEEE Intl Conf. 1999:969-973.
- [67] Lim S., Lu G. J. Spatial statistics for content based image retrieval[C]. International Conference on Information Technology: Computers and Communications California, IEEE Computer Society, 2003:28-30.
- [68] 汪慧兰, 赵海峰, 罗斌. 基于局部颜色空间特征的图像检索[J]. 计算机技术与发展, 006, 16(1):76-78.
- [69] 王涛, 胡事民, 孙家广. 基于颜色—空间特征的图像检索[J]. 软件学报, 2002, 13(10): 2031 -2036.
- [70] Hsu W., Chua T. S., Pung H. K. An integrated color-spatial approach to content-based image retrieval[C]. In: Proc ACM Multimedia'95 Conference. San Francisco, 1995, 305-313.
- [71] Dimai A. Spatial encoding using differences of global features[C]. In Proc. of SPIE storage and retrieval for image and video database IV, 1997, 3022: 352-360.
- [72] 牛蕾, 倪林, 苗原. 基于颜色空间分布特征的图像检索[J]. 计算机工程与应用. 2005, 4: 50 - 53.
- [73] 张志安, 骆斌. 一种新的颜色空间特征图像检索[J]. 西安电子科技大学学报(自然科学版), 2008,35(3):577-580.
- [74] Rao A., Srihari R., Zhang Z. Spatial Color Histogram for Content-Based Retrieval[C]. Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence, 1999:183-186.
- [75] 安志勇. 基于内容的图像检索关键技术研究[D]. 西安: 西安电子科技大学, 2008.



- [76] 贾元元, 基于形状的图像数据库检索技术研究[D]. 西安: 西安电子科技大学, 2004.
- [77] Hu M. K. Visual pattern recognition by moment invariants[J]. IRE Transactions on Information Theory, 1962, 8(2) : 179- 187.
- [78] M. R. Teague. Image analysis via the general theory of moments[J]. Journal of the Optical Society of America, 1979, 70(8):920-930.
- [79] 杨冠羽, 舒华忠, 周卫平. 一种新的灰度图像 Legendre 矩的快速算法[J]. 计算机学报, 2004, 27(12):1642-1647.
- [80] 马丽红, 张学东, 赵忠夫. 基于正交 Legendre 矩的虹膜识别算法[J]. 计算机应用, 2006, 26(4): 787-789.
- [81] Alreza Khotanzad, Yaw Hua Hong. Invariant Image Recognition by Zernike Moments[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1990, 12(5):489-497.
- [82] 李海啸, 姜璐, 舒华忠. 基于 Zernike 矩和 BP 神经网络的纹理分割[J]. 东南大学学报 (自然科学版), 2005, 35(3):199-201.
- [83] 王金岩, 芮挺, 丁健. 基于 Zernike 矩和 SVM 的不变性目标识别[J]. 电光与控制, 2008, 15(11):1-4.
- [84] 郭丹, 闫德勤, 吴晓婷. 一种基于 Zernike 矩形状检索的新算法[J]. 计算机科学, 2010, 37(11):
- [85] C.H.Teh, R.T.Chin. On Image Analysis by the Method of Moments[J]. IEEE Trans. Pattern Anal. Mach. Intelligence. 1988, 10(4):496-513.
- [86] 王勃飞. 基于 Pseudo-Zernike 矩的数字识别研究[D]. 武汉: 华中科技大学, 2007.
- [87] 夏婷, 周卫平, 李松毅等. 一种新的 Pseudo-Zernike 矩的快速算法[J]. 电子学报, 2005, 33 (7): 1295-1298
- [88] 陈平, 李垒, 任越美. 基于区域伪 Zernike 矩的商标图像检索[J]. 计算机仿真, 2010, 27(8): 270-272.



- [89] Mukundan R, Ong S H, Lee P A. Image Analysis by Tchebichef Moments[J]. IEEE Trans. on Image Processing, 2001, 10(9): 1357-1364.
- [90] 赵杰, 王晖, 何冰. 基于 Tchebichef 矩和小波提升的数字水印算法[J]. 计算机工程, 2009, 35(11): 113-115.
- [91] M. Emre Celebi, Y. Alp Aslandogan. A Comparative Study of Three Moment-Based Shape Descriptors. Proceedings of the International Conference on Information Technology: Coding and Computing, 2005, 1: 788 -793.
- [92] Pew-ThianYap, Raveendran. Image Analysis by Krawtchouk Moments[J]. IEEE Trans. On Image Processing, 2003, 12(11): 1367-1376.
- [93] A. Sajjanhar, J. Wright. An experimental study of moment invariants and Fourier descriptors for shape based image retrieval[C]. Proceedings of the Second Australia Document Computing Symposium, Australia, 1997:46-54.
- [94] 姚玉荣, 章毓晋. 利用小波和矩进行基于形状的图像检索[J]. 中国图像图形学报, 2000, 5(3):206-210.
- [95] Mallat S, Hwang W L. Singularity detection and processing with wavelets[J]. IEEE Trans Info Theory, 1992, 38(2):617-643.
- [96] Li Junhong, Pan Quan, Zhang Hongcai, *et al.* Image recognition using Radon transform[C]. International Conference on Intelligent Transportation Systems IV: Image Analysis, Shanghai, China: IEEE Computer Society, 2003, 4: 741-744.
- [97] Osama K. Al-Shaykh, John F. Doherty. Invariant Image Analysis Based on Radon Transform and SVD[J]. IEEE Transactions on Circuits and Systems-11: Analog and Digital Signal Processing, 1996,



- 43(2):123-133.
- [98] 韩军伟. 基于内容的图像检索技术研究[D]. 西安: 西北工业大学, 2003.
- [99] 曾智勇. 基于内容图像数据库检索中的关键技术研究[D]. 西安: 西安电子科技大学, 2006.
- [100] Gross M. H., Koch R., Lippert L., Dreger A. Multiscale image texture analysis in wavelet spaces[C]. In Proc. IEEE Int. Conf. on Image Proc., 1994.
- [101] Niblack. The QBIC project: querying images by content using color, texture, and shape[C]. Proc. of SPIE, Storage and Retrieval for Image and Video Databases, 1993, 1908: 173-187.
- [102] Gotlieb C. C., Kreyszig H. E. Texture descriptors based on co-occurrence matrices[J]. Comput. Vis., Graphics, and Image Proc. 1990, 51:70-86.
- [103] T.Ojala, M.Pietikainen, David Harwood. A comparative study of texture measures with classification based on feature distributions[J]. Pattern recognition, 1996, 29:51-59.
- [104] 蔡蕾. 基于纹理的图像检索方法研究[D]. 长春: 吉林大学, 2005.
- [105] Timo Ojala, Matti Pietikainen. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns[J]. IEEE Transactions on pattern analysis and machine intelligence, 2002, 24(7):971-987.
- [106] Laws K L. Rapid texture identification[C]. In Proceedings of SPIE, 1980, 238: 376-380.
- [107] 韦娜. 基于纹理特征的图像检索技术研究 with 实现[J]. 西安: 西北大学, 2003.
- [108] 盛文, 柳键. 图像纹理分析方法及其最新进展[J]. 无线电工程,



- 1998, 128 (5) : 8-14.
- [109] Manjunath B. S., Ma W. Y. Texture features for browsing and retrieval of image data[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1996, 18(8):837-842.
- [110] 褚标. 小波理论在图像去噪与纹理分析中的应用研究[D]. 合肥: 合肥工业大学, 2008.
- [111] A.Grossman, J.Morlet. Decomposition of hardy functions into square Integrable wavelets of constant shape[J]. SIAM J.Math.Anal.1984, 15(4):723-736.
- [112] Daubechies. Orthonormal bases of compactly supported wavelets.[J] Comm. On Pure and Appl. Math., 1988, 41(7): 909-996.
- [113] Mallat S. Multiresolution approximation and wavelet orthonormal bases of $L^2(R)$ [J]. IEEE Trans.Math. Soc. 1989.9, 315: 69-87.
- [114] Mandal M. K., Aboulnasr T. Fast wavelet histogram techniques for image indexing[J]. Computer Vision and Image Understanding, 1999, 75(1/2): 99-110.
- [115] 汪祖媛, 庄镇泉, 何劲松. 基于形状的小波变换系数广义高斯分布图像检索算法[J]. 电子学报, 2003,31(5):765-768.
- [116] 尚赵伟, 张明新, 沈钧毅. 基于 KL 距离和双密度小波变换的纹理图像检索[J]. 微电子学与计算机, 2006,23(2):13-19.
- [117] JuHan, Kai-KuangMa. Rotation-invariant and scale-invariant Gabor features for texture image retrieval[J]. Image and Vision Computing, 2007 (25):1474-1481.
- [118] H. Abrishami Moghaddam, M. Saadatmand-Tarzjan. Gabor Wavelet Correlogram Algorithm for Image Indexing and Retrieval[C]. Proc. 18th International Conference on Pattern Recognition, ICPR 2006, 2: 925-928.



- [119] Strela V, Heller P N, Strang G. The application of multiwavelet filter-banks to image processing[J]. IEEE Trans on Signal Processing, 1999;8(4):548-563.
- [120] 尚赵伟, 刘贵忠, 赵平. 基于多小波直方图的纹理图像检索[J]. 西安交通大学学报, 2005, 39(2):123-125.
- [121] 魏海, 沈兰荪. 小波变换域内基于方向梯度相角直方图的图像检索算法[J]. 电路与系统学报, 2001,6(2):20-24.
- [122] 张国云, 彭仕玉. 基于反对称双正交小波的图像压缩算法[J]. 计算机工程, 2008,34(6):205-207.
- [123] 张德胜, 罗晓辉, 张遵伟. 基于双树复小波和灰度共生矩阵的纹理图像检索[J]. 西华大学学报(自然科学版), 2010,29(3):87-90.
- [124] Nick Kingsbury. The Dual tree Complex Wavelet Transform: a New Technique for Shift Invariance and Directional Filters[C]. Proc .8th IEEE DSP Workshop, Bryce Canyon, 1998.
- [125] Selesnick I W. The double-density dual-tree DWT[J]. IEEE Trans. on Acoustics, Speech, and Signal Processing, 2004, 52(5): 1304-1314.
- [126] 尚赵伟, 张明新, 沈钧毅等. 基于双密度小波变换的纹理图像检索[J]. 西安交通大学学报, 2005, 39 (10):1081-1084.
- [127] 郭巍, 张平, 陈曦. 孔径雷达图像降噪研究基于双密度双树复数小波变换的合成[J]. 电子学报, 2009,37(12):2747-2652.
- [128] 李鹏, 喻罡, 冀晓燕. 基于双密度双树小波变换的超声图像降噪[J].系统仿真学报, 2009,19(12):5797-5801.
- [129] Pun C M, Lee M C. Rotation invariant texture feature for content based image retrieval[J]. Computer Vision and Image Understanding, 2003, 89(1): 24-43.
- [130] Pun C M, Lee M C. Log-Polar wavelet energy signatures for rotation and scale invariant texture classification[J]. IEEE Transactions on



- Pattern Analysis and Machine Intelligence, 2003, 25(5):590-603.
- [131] Kourosh J.K, Hamid S.Z. Rotation-invariant multiresolution texture analysis using Radon and wavelet transform[J]. IEEE Trans.Image Process, 2005, 14(6): 783-794.
- [132] 安志勇, 曾智勇, 崔江涛. 基于 Radon 和小波变换的纹理图像检索[J].北京邮电大学学报, 2007,30(5):5-8.
- [133] Xiong H L, Zhang T X, Moon Y. S. A translation- and scale-invariant adaptive wavelet transform[J].IEEE Trans. Image Processing, 2000, 9(12): 2100-2108.
- [134] 潘泓, 夏良正. 具有平移、尺度和旋转不变的小波变换[J]. 信号处理, 2004, 20(2):147-151
- [135] L. Itti, C. Koch, E. Niebur. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Trans. On PAMI, 1998, 20(11): 1254-1259.
- [136] Y.N. Deng, C. Kenney, M.S. Moore. Peer group filtering perceptual color image quantization[C]. Proc. IEEE Intl. Symposium on Circuits Systems, 1999, 4: 21-24.
- [137] A. Gersho and R.M. Gray. Vector quantization and signal compression[M]. Kluwer Academic, Norwell, MA, 1992.
- [138] H.D. Cheng, Y.H. Chen, X.H. Jiang. Thresholding using two-dimensional histogram and fuzzy entropy principle[J]. IEEE Trans. On Image Processing, 2000, 9:732-735.
- [139] 王向阳, 杨红颖, 胡峰丽. 基于感兴趣区的小波域彩色图像检索新方法[J]. 中国图像图形学报, 2006,11(2):175-179.
- [140] 陆伟, 倪林. 利用颜色和熵提取感兴趣区域的感性图像检索[J]. 中国图像图形学报, 2006,11(4):492-497.
- [141] 曾智勇, 安志勇, 周利华. 显著兴趣点颜色及空间分布熵的图像



- 检索方法[J]. 红外技术, 2007, 29(3):160-163.
- [142] 徐斌. 兴趣点及灰度差分不变量在图像匹配中的应用研究[D]. 西安: 西安电子科技大学, 2001.
- [143] Deriche L. S., Nagel H. N. On the selection critical points and local curvature extrema of region boundaries for interframe matching[C]. International conference on Pattern Recognition, 1982, 542-545.
- [144] Hua H. Z., Qian Y. A direct corner detecting algorithm[C]. International conference on Pattern Recognition, Paris, 1986, 853-856.
- [145] Rangarajan K., Shah M. Brackley D. V. Optimal corner detector[J]. Comput. Vis Graph, Image Processing, 1989, 48:230-245.
- [146] Deriche R., Blaszkowski T. Recovering and characterizing image features using an efficient model based approach[C]. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1993, 530-535.
- [147] Moravec H. P. Towards automatic visual obstacle avoidance[C]. In: Proceedings of the 5th International Joint Conference on Artificial Intelligence. Cambridge, USA: MIT Press, 1977, 584-599.
- [148] Harris C., Stephens M. A Combined Corner and Edge Detection[J]. Image Vision Computing, 1998, 6: 121-128.
- [149] Kitchen L., Rosenfeld A. Gray-Level Corner Detection[J]. Pattern Recognition Letters, 1982, 1: 95-102.
- [150] Lee. J. S., Sun Y. N., Chen C. H. Multiscale corner detection by using wavelet transform[J]. IEEE Trans. on Image Processing, 1995, 4(1):100-104.
- [151] Mallat S. A theory for multiresolution signal decomposition: the wavelet representation[J]. IEEE Trans. on PAMI, 1989,



- 11(7):674-693.
- [152] 孟繁杰, 郭宝龙. 使用兴趣点局部分布特征及多示例学习的图像检索方法[J]. 西安电子科技大学学报(自然科学版), 2011, 38(2):47-53.
- [153] Zhang Q, Goldman S A. EMDD: an Improved Multiple instance Learning Technique[C]. Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2002:1073-1080.
- [154] Candes E J. Ridgelets: Theory and Applications[D]. Ph.D. dissertation. Dept. Statistics, Stanford Univ., 1998.
- [155] Donoho D L, Duncan M R. Digital curvelet transform: strategy, implementation and experiments[C]. Proc.SPIE, 2000, 4056:12-29.
- [156] Minh N. Do, M. Vetterli. The contourlet transform: an efficient directional multiresolution image representation[J]. IEEE Transactions on Image Processing, 2003, 14(12):2091-2106.
- [157] Velisavljevic V, Beferull-Lozano B, Vetterli M, et al. Directionlets: anisotropic multi-directional representation with separable filtering[J]. IEEE Transactions on Image Processing, 2006, 15(7): 1916- 1933.
- [158] 侯彪, 刘佩, 焦李成等. 基于改进 Wedgelet 变换的 SAR 图像边缘检测[J]. 红外与毫米波学报, 2009, 28(7): 396-400.
- [159] R. Eslami and H. Radha. Wavelet-based contourlet transform and its application to image coding[J]. IEEE International Conference on Image Processing, 2004, 5: 3189-3192.
- [160] P J. Burt, E. H. Adelson. The Laplacian as a compact image code[J]. IEEE Transactions on Communication, 1983, 31(4): 532-540.
- [161] R. H. Bamberger, M. J. T. Smith. A filter bank for the directional decomposition of images: theory and design[J]. IEEE Transactions on Signal Processing, 1992, 40(4): 882-893.



-
- [162] 黄传波, 邵杰, 万鸣华等. Contourlet 变换在图像检索中的应用[J]. 计算机工程与应用, 2009 45(3): 24-27.
- [163] 杨家红, 许灿辉, 王耀南. 基于 Contourlet 广义高斯模型的纹理图像检索[J]. 中国图像图形学报, 2007, 12(4): 691-694.
- [164] Ch.Srinivasa rao, S.Srinivas kumar, B. N.Chatterji. Content based image retrieval using contourlet transform[J]. ICGST-GVIP Journal, 2007, 7(3): 9-15.
- [165] Arthur L. Cunha, Jiangping Zhou, Minh N. Do. The Nonsubsampled contourlet transform: theory, design, and applications[J]. IEEE Trans Image Processing, 2006, 15(10): 3089-3101.
- [166] 梁昕. NSCT 变换在 SAR 图像检索中的应用与实现[D]. 南京: 南京航空航天大学, 2008.
- [167] 连学强. 无下采样 Contourlet 变换及其在图像处理中的应用研究 [D]. 厦门: 厦门大学, 2008.
- [168] 练秋生, 孔令富. 圆对称轮廓波构造[J]. 计算机学报, 2006, 29 (4) : 652-657.
- [169] 练秋生, 李芹, 孔令富等. 融合圆对称轮廓波统计特征和 LBP 的纹理图像检索[J]. 计算机学报, 2007, 30(12): 2198-2204.
- [170] 陈新武, 刘玮, 田金文. 复数轮廓波变换纹理图像检索系统[J]. 光电工程, 2009, 36(2): 111-115.
- [171] Chen D, Li Q. The Use of Complex Contourlet Transform on Fusion Scheme[C]. Proceedings of World Academy Of Science, Engineering And Technology 2005, 7: 342-347.
- [172] DAI Shao-wei, SUN Yan-kui, TIAN Xiao-lin, et al. Image Denoising Based on Complex Contourlet Transform[C]. International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China, Nov 2-4, 2007. Washington D C, USA: IEEE, 2007, 4: 1742-1747.
-

《基于视觉内容的图像检索技术研究》读者调查表

尊敬的读者：

欢迎您参加读者调查活动，对我们的图书提出真诚的意见，您的建议将是我们创造精品的动力源泉。

为方便大家，我们提供了两种填写调查表的方式：

1. 您可以登录 <http://yydz.phei.com.cn>，进入“读者调查表”栏目，下载并填好本调查表后反馈给我们。
2. 您可以填写下表后寄给我们（北京海淀区万寿路 173 信箱电子技术分社 邮编：100036）。

姓名：_____ 性别：☐ 男 ☐ 女 年龄：_____ 职业：_____
电话：_____ 移动电话：_____
传真：_____ E-mail：_____
邮编：_____ 通信地址：_____

1. 影响您购买本书的因素（可多选）：

☐封面、封底 ☐价格 ☐内容简介 ☐前言和目录 ☐正文内容
☐出版物名声 ☐作者名声 ☐书评广告 ☐其他_____

2. 您对本书的满意度：

从技术角度	<input type="checkbox"/> 很满意	<input type="checkbox"/> 比较满意	<input type="checkbox"/> 一般	<input type="checkbox"/> 较不满意	<input type="checkbox"/> 不满意
从文字角度	<input type="checkbox"/> 很满意	<input type="checkbox"/> 比较满意	<input type="checkbox"/> 一般	<input type="checkbox"/> 较不满意	<input type="checkbox"/> 不满意
从版式角度	<input type="checkbox"/> 很满意	<input type="checkbox"/> 比较满意	<input type="checkbox"/> 一般	<input type="checkbox"/> 较不满意	<input type="checkbox"/> 不满意
从封面角度	<input type="checkbox"/> 很满意	<input type="checkbox"/> 比较满意	<input type="checkbox"/> 一般	<input type="checkbox"/> 较不满意	<input type="checkbox"/> 不满意

3. 您最喜欢书中的哪篇（或章、节）？请说明理由。

4. 您最不喜欢书中的哪篇（或章、节）？请说明理由。

5. 您希望本书在哪些方面进行改进？

6. 您感兴趣或希望增加的图书选题有：

邮寄地址：北京市万寿路 173 信箱电子技术出版分社 张剑 收 邮编：100036
电 话：（010）88254450 E-mail: zhang@phei.com.cn